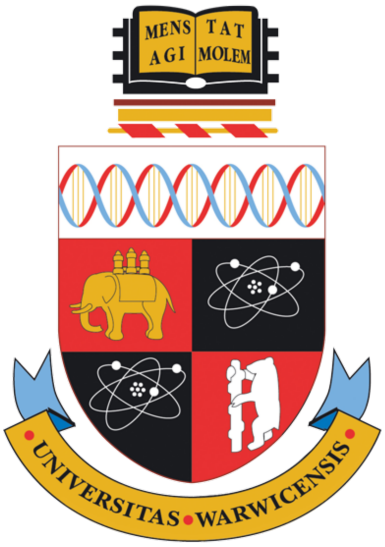

Studies of the rare decays $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ with the LHCb experiment

by

Edward James Millard

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

December 2020

Contents

“ *Knowledge is an unending adventure at the edge of uncertainty* ”

Jacob Bronowski,

Acknowledgments

Firstly, I would like to thank my supervisor Dr. Tom Blake for his support over the last four years. Working with him has been a delight and it was the combination of his first-rate physics knowledge and endless patience which have allowed this thesis to see the light of day. Secondly, I would like to say a huge thank you to Dr. Tom Latham for his help on all matters ranging from programming bugs and Ganga meltdowns to advice on which restaurants to try in San Francisco. Thank you also Prof. Tim Gershon, Dr. Michal Kreps and Dr. Steve Boyd who all inspired me to pursue further study in particle physics when I was an undergraduate.

I would like to thank my fellow PhD students and office members, past and present, for the camaraderie over the past few years. Thanks to Dr. Andrew Crocombe for the initial tutorials and thanks to Dr. Abhijit Mathad, Arnau Gonzalo, Tom Lord and Dr. Cayo Sobral for the many late night discussions where we tried to wrap our heads around various problems - physics related and otherwise. Thanks also to Dr. Niladri Sahoo, Tom Jones, Bhagyashree Pagare, Ross Hunter, Andy Morris and Flavia Cicala for making office 449 a delightful place to work in. A particular thank you must go to my dear friends James Pillow and Sammy Valder as we have been a unit since day one. I look forward to our next frame of pool.

A very special thank you to my family, Mum, Dad and Dan for their love and support over the PhD and always. Also thanks to my girlfriend, Valerie, for being by my side since Heronbank and for always telling me - 加油!

Finally, I would like to thank Prof. Stan Brodsky for the pickaxe idea. Unfortunately, this will have to be put on hold.

Edward Millard

Declarations

The work presented in this thesis is all of my own work, unless it is specifically referenced to the contrary. This thesis has not been submitted, in any form, to this or any other university for another qualification.

Edward Millard

December 22, 2020

Abstract

Flavour changing neutral currents only occur at loop level within the Standard Model of particle physics. This makes these processes extremely rare and potentially sensitive to interference with currently undiscovered physics that lies beyond the scope of the Standard Model. To date, tensions with the Standard Model have been found in the studies of $b \to sl^+l^-$ processes at large hadronic recoil. This thesis presents a branching fraction analysis of the $b \to d(s)\mu^+\mu^-$ processes $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ using the data recorded by the LHCb collaboration during Run 1 and Run 2 of the Large Hadron Collider. The presence of the $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ decays are confirmed with a statistical significance in excess of 12 standard deviations. The branching fraction of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ process is measured to be $(2.35 \pm 0.26 \pm 0.18) \times 10^{-8}$ while the branching fraction of the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ process is measured to be $(5.35 \pm 0.64 \pm 0.24) \times 10^{-8}$ where the first uncertainties are statistical and the second systematic. In addition, this thesis presents for the first time the branching fraction measurements for the decay modes in two separate regions of the invariant mass of the dimuon pair, q^2. The B^0 decay is observed for the first time at both large and small hadronic recoil while the B_s^0 decay is observed for the first time at large hadronic recoil.

CHAPTER 1

Introduction

> *There will come a time when our descendants will be amazed that we did not know things that are so plain to them*

Seneca,

It is a common misconception that the idea that matter is made up of particles was first conceived of within the twentieth century. Long before the birth of what we now call the scientific method, natural philosophers the world over pondered the mysteries of nature and expressed interest in the building blocks of the material world. Historical records tell us that Jainists in ancient India, from the 9th to the 5th century BCE, referred to the *paramanu*, which were tiny, indivisible particles which made up the "non-living" part of the universe [1]. Furthermore, in ancient Greece in the 5th century BCE, the doctrine of atomism was furthered by the philosopher Democritus who reasoned that it was unthinkable to believe that matter could be divided ad infinitum [2]. [1] Of course, these early ideas were philosophical in nature and completely untestable due to the limitations of technology at the time. As a result, many of these musings from the ancient world were effectively forgotten about at the end of the classical period with only a few, such as intellectual factions of the Islamic Golden Age, keeping them alive [3].

Centuries later, new scholars of the Western enlightenment returned to these concepts and further developed them. Robert Boyle [4], of thermodynamics fame, supposed that all of the chemical elements were composed of particles of various sizes. Sir Isaac Newton concluded, like the Islamic scholars before him, that light may be composed of tiny "corpuscles" and John Dalton

[1] The word atom comes from the Greek word *atomos*, meaning "uncuttable."

built on the Democritic idea by pushing forward his atomic theory [5]. It can be argued that interest in the field truly began to accelerate at the end of the nineteenth century with the discovery of the electron in 1897 by J. J. Thomson which moved the discipline from the atomic to the sub-atomic for the first time [6]. Indeed, following on from Thomson's discovery of the electron, discoveries came thick and fast. Soon, another key figure, Ernest Rutherford, demonstrated that the atom was mainly comprised of empty space before he went on, in 1918, to discover the existence of the proton [7] which was then complemented by James Chadwick's discovery of the neutron in 1932 [8].

As the twentieth century got underway, particle physics entered a golden age -- primarily driven by the advent of quantum mechanics, special relativity and the subsequent revolutions of modern physics. The recognition of this new set of theories led to the development of modern chemistry via the Bohr model of the atom, the unlocking of the secrets of both nuclear fission and fusion and new insights into the nature of magnetism [9] -- a phenomenon known about for thousands of years but which had never been well understood.

Following on from the discovery of the muon in 1937 [10] more exotic particles, those not found easily in nature, were hypothesised. Simultaneously, thanks to rapid developments in accelerator and detector technology, higher energies could be probed which led to the prediction, discovery and classification of a plethora of new particles and the discovery of an underlying, three generational structure to matter. Within the second half of the century the rapidly maturing set of theories were unified into the modern Standard Model of particle physics [11] which has boasted numerous successful predictions and which culminated, famously, with the discovery of the Higgs boson in 2012 [12, 13].

Today, the study of particle physics has reached an inflection point; despite the Standard Model's successes the model is known to be incomplete and a superior understanding of the universe is currently lacking. There are currently many different areas of study within the field, each of which is being experimentally probed to find a statistically significant derivation from the Standard Model which leads to the next big discovery. This thesis focuses on the heavy flavour sector which concerns transitions between the different types of heavy quark within the Standard Model. Within this sector, international collaborations such as the LHCb experiment are currently observing some tensions with the established model -- potentially showing the first signs of new physical phenomena.

One area of current experimental focus in heavy flavour physics is the study of the rare decays of beauty hadrons. This thesis, using data from the LHCb experiment, presents an analysis of the rare decays of the beauty hadrons B^0 and B_s^0 to a $\pi^+\pi^-\mu^+\mu^-$ final state.

The organisation of this thesis is as follows. Chapter 2 discusses the Standard Model of particle physics, provides a motivation for the study of rare decays within heavy flavour physics,

and reviews some previous experimental results. Chapter 3 describes the functionality of both the Large Hadron Collider and the LHCb detector. Chapter 4 introduces the analysis strategy while Chapters 5, 6, 7, 8 and 9 discuss in detail the various steps of the analysis itself. The results are analysed within the context of the wider literature in Chapter 10. Finally, additional information regarding the LHCb data flow and the service task completed for the collaboration is included in Appendix A and some fit parameters, not included within the main body of Chapter 8, are included in Appendix B. Some additional figures are included in Appendix C for reference.

CHAPTER 2

Theoretical Background

> *He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast*
>
> Leonardo Da Vinci,

Within this chapter the theory motivating the thesis work is presented. This includes a general introduction to the Standard Model of particle physics with a particular emphasis on flavour physics and b-hadron decays. In addition, some of the current exciting flavour anomalies as measured by both the B-factories and LHCb are discussed. The anomalies include recent branching fraction measurements, angular analyses and lepton flavour universality measurements which all show tensions with the Standard Model within the same kinematic region. A motivation for the thesis work, in the context of these anomalies, is presented at the end of the chapter.

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a relativistic quantum field theory which describes the fundamental constituents of the universe and the dynamics that govern them [15–17]. Within the model all of the known particles are understood to be excited quanta of the fundamental fields which permeate the universe and these particles are classified by their quantum numbers, as visualised in Figure 2.1.

Particles with a half integer spin are known as fermions and are divided into the quarks, of which there are six (up, down, charm, strange, top and bottom) and the leptons of which three are

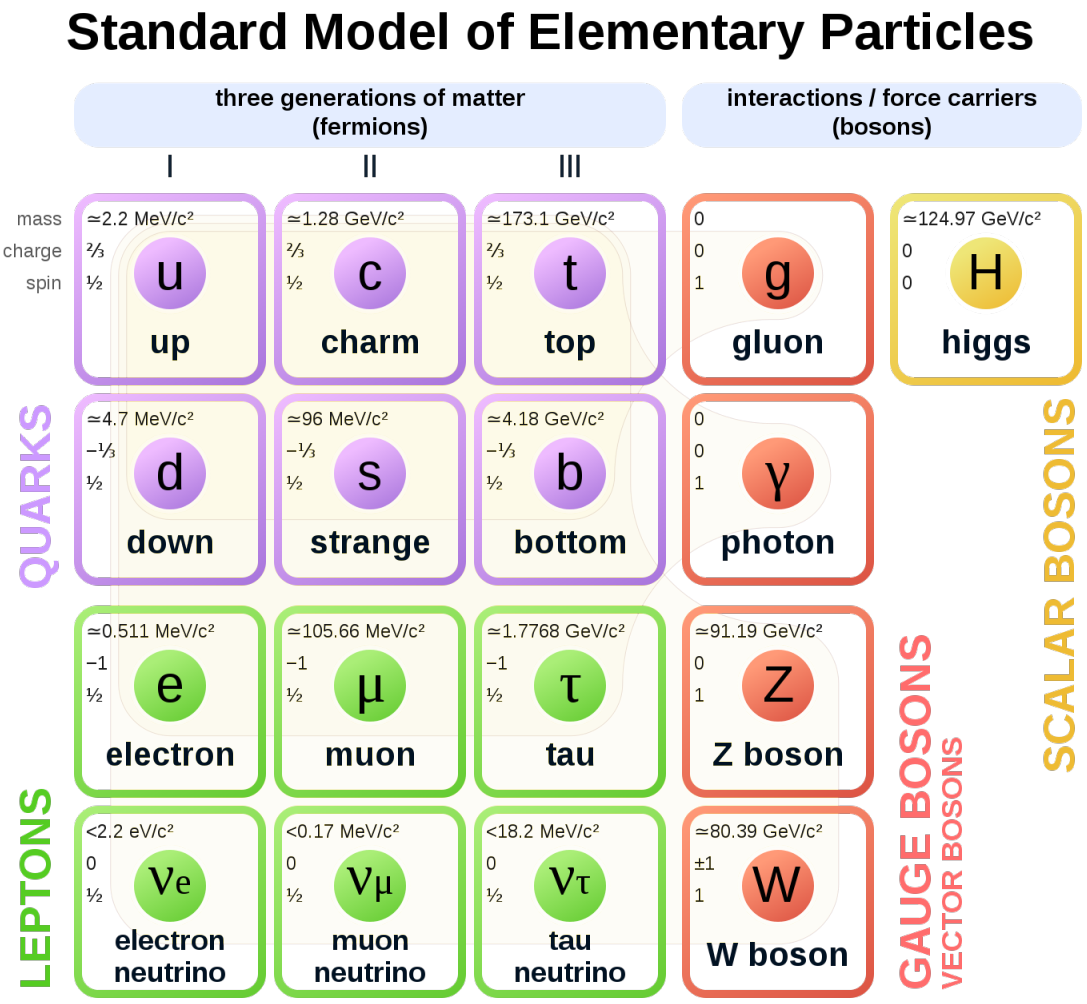

Figure 2.1: The constituents of the Standard Model of particle physics [14].

charged (electron, muon and tau) and three are neutral (electron neutrino, muon neutrino and tau neutrino) [18]. These different types, or *flavours*, of fermion are arranged into three *generations* of increasing mass whereby fermions in either the second or third generation eventually decay back to the first due to Nature's proclivity to maintain the universe in the lowest possible energy state [19, 20]. It is for this reason that the atom is comprised of up and down type quarks with orbiting electrons leaving the more exotic fermions within the SM absent from daily life.

Alongside these fermions are their antimatter equivalents; partners to the fermions which share the same mass but have oppositely-signed quantum numbers meaning that an interaction be-

tween a matter particle and it's antimatter counterpart results in annihilation. This set of fermions then interacts with the fundamental forces of nature via the full-integer spin gauge bosons. The strong nuclear force is mediated by the gluon as described by quantum-chromodynamics (QCD), the electromagnetic interaction is mediated by the photon as described by quantum-electrodynamics (QED) and the weak nuclear force is mediated by the Z^0 and $W^{\pm}$ bosons with the underlying theory being merged with electromagnetism as part of the overarching electroweak theory [21].

Whether or not a fermion will interact via each force depends on whether the fermion possesses the corresponding hypercharge. For example, only fermions with a non-zero electrical charge may interact electromagnetically. The charge equivalent for the weak interaction is the weak hypercharge, possessed by all fermions, while for the strong interaction it is the colour charge which, uniquely comes in three varieties, red, green and blue. Furthermore, the non-abelian symmetry structure of QCD mandates that objects bound together by the strong force must be colour singlets. As a result, the quarks, which also possess colour charge, are unable to exist independently and therefore are bound together into *hadrons* [22]. There are, to date, several observed combinations of quarks and anti-quarks under the umbrella of hadrons including the $q\bar{q}$ meson, the qqq baryon, and, as recently confirmed to exist by the B-factory experiments and the LHCb experiment, the exotic $qq\bar{q}\bar{q}$ tetraquark and $qqq\bar{q}\bar{q}$ pentaquark [23–31]. The binding together of the quarks in various combinations leads to an extensive list of hadrons and the so-called "particle zoo" as was predicted and classified by Gell-Mann and Zweig in the 1960s as part of their Quark Model [32, 33].

Finally, alongside the quarks, leptons and gauge bosons, the SM includes a single scalar boson, the now famous Higgs boson. This particle exists as an excitation of a field which, through the Higgs mechanism (discussed in more detail in Section 2.4), broke the electroweak symmetry within the early universe resulting in most SM particles gaining a mass.

A noticeable absence from the SM is a gauge boson responsible for mediating the force of gravity although such a boson, normally referred to as the Graviton, has been hypothesised. Furthermore, gravity is the only fundamental force in physics (that we currently know of) which has not been incorporated into the SM and all attempts to combine Einstein's theory of General Relativity with quantum mechanics have, to date, been fruitless [34].

2.2 The Insufficiency of the Standard Model

The successes of the SM are hard to understate and they culminated in 2012 with the discovery of the Higgs boson by the ATLAS and CMS collaborations at CERN [12, 13]. Indeed, the SM of particle physics can lay claim to the most accurately validated theory in scientific history with the

measurement of the anomalous magnetic moment of the electron [35]. However, despite these successes, the SM does not address some of the phenomena we witness in nature and therefore cannot be considered a complete picture.

Aside from the inability of the SM to incorporate gravity, as mentioned above, several other key phenomena remain unexplained such as the nature of dark matter and dark energy. Together these make up $\sim$95% of the energy density of the universe and their existence is inferred from astrophysical data such as galactic rotation curves, gravitational lensing and the cosmic microwave background [36, 37].

In addition, the observed matter-antimatter asymmetry seen in the universe is not understood. It is assumed that in the early, hot, universe there was a statistical equilibrium between baryonic and anti-baryonic matter which evidently is no longer the case. If the universe indeed did once possess this equilibrium then Nature, in order to have evolved away from this initial state, must possess the ability to fulfill the three so-called *Sakharov* conditions that explain under what conditions matter and antimatter can be produced at different rates [38]. The conditions are:

1. A mechanism to violate the conservation of Baryon number must exist;
2. Interactions must have taken place out of thermal equilibrium in order to allow the state of the universe to evolve;
3. There must be mechanisms which violate both C and CP symmetry.

Mechanisms which violate CP symmetry are known of in the SM (discussed further in Section 2.3) however, they do not account for the size of the CP violating effect required to explain the state of the modern universe [39].

Additional conundrums in particle physics exist which the SM also fails to explain. These include:

- We do not know why there are three generations of fermions;
- QCD results in a CP violating term yet no strong CP violation is observed experimentally [40];
- The Higgs mass (at 125 GeV) seems very low given existing large quantum loop contributions [41];
- There seem to be many (19) free parameters in the SM -- why is this?

How to go about finding an explanation to these big questions is a matter of debate within the field of particle physics. There are theories which address some of these problems by invoking the

existence of new, heavy, particles which only exist at the TeV energy scale [42]. Currently, and as is the focus of this thesis, there are some exciting tensions currently seen experimentally in the rare decays of beauty hadrons (see Section 2.10). Although these tensions are yet to reach a 5 σ confirmation level they may point to a problem with the SM and, potentially, to physics beyond the scope of the SM which, one would hope, may start to shed some light on these conundrums.

2.3 Symmetries of the Standard Model

The role of symmetries within particle physics is a crucial one. In 1917 Emmy Noether published her theorem which highlighted the relationship between symmetries and the dynamics of a system [43]. Her theorem states that every symmetry of nature yields a corresponding conservation law. Some examples include the symmetry of a system under a translation in time corresponding to the conservation of energy, and the symmetry of the electromagnetic interaction under a local gauge transformation which results in the conservation of electric charge.

The SM exhibits invariance under continuous local gauge symmetry transformations which belong to the

$$SU(3)_C \times SU(2)_L \times U(1)_Y \tag{2.1}$$

group [44]. Here, $SU(3)_C$ represents QCD with C a conserved colour charge, while $SU(2)_L \times U(1)_Y$ represents the symmetry of the electroweak theory with L denoting the left-handed chirality of the weak interaction and Y denoting conserved weak hypercharge. This symmetry structure is not obeyed at all energy scales however. Indeed, in today's cooled universe the electroweak symmetry has been broken by the Higgs mechanism. Also, remembering that the SM is only an effective, low-energy theory, a high-energy grand unified theory most likely will have a different symmetry structure which was broken very quickly within the early universe [45, 46].

In addition to the continuous form of symmetry, discrete symmetries are also present, such as symmetries of space-time. At various points in the development of the SM certain discrete symmetries have been considered to be fundamental symmetries of the universe. These discrete symmetries are:

- ***C*** - Charge Conjugation: The symmetry of the system when replacing a particle with it's anti-particle by inverting the sign of all quantum numbers;
- ***P*** - Parity: The symmetry of the system when inverting the sign of all spatial space-time co-ordinates;
- ***T*** - Time: The symmetry of the system when inverting the temporal co-ordinate of space-time.

Today, it is known that these three symmetries are not good symmetries of the universe when applied individually, i.e. transformation under these symmetries alone does not guarantee the invariance of a system. Chronologically, *P* symmetry was long considered a good symmetry and indeed is conserved in both electromagnetic and strong interactions. However, in 1957 an experiment was performed by Wu showing that the direction of electrons emitted from the decay of ^{60}Co atoms was completely asymmetric to their spin orientation, thereby demonstrating that the weak force was violating parity [47]. At the time the natural assumption of particle physicists was that the application of the combination of the operators *C* and *P*, CP, must be invariant in nature. This too was shown to not be the case in 1964 by Cronin and Fitch in their experiment where they found evidence of the CP violating decay $K_L^0 \to \pi^+\pi^-$ [48] where a CP odd $K_{\rm L}^0$ decays to a CP even $\pi^+\pi^-$ pair.

Given it's importance in explaining the matter-antimatter asymmetry of the universe, much experimental effort has been expended trying to probe the extent of CP violation. As of 2020 CP violation has been observed independently in K, B^0, B_s^0 and (recently) D mesons [49–52]. In addition, there are searches ongoing for CP violation in the baryon sector [53–55] and there are even hints that violation may be present in the neutrino sector as part of the neutrino mixing model with experimental confirmation yet to be achieved [56].

Lastly and continuing the pattern, the *T* symmetry has been observed to be violated in an analysis performed by the BABAR experiment at SLAC which measured T violating parameters in neutral B meson decays [57]. However, despite the confirmation of both T and CP symmetry breaking, the application of *C*, *P* and *T*, known as CPT, is, according to the CPT theorem, a good symmetry of nature [58, 59]. The theorem states that CPT invariance is upheld in a relativistic quantum field theory as long as Lorentz invariance is respected, a condition which the SM fulfills. To date, no experiment has witnessed any evidence of CPT violation.

2.4 The Higgs Mechanism and Spontaneous Symmetry Breaking

The symmetries of nature result in some strange and conceptually challenging phenomena. For example, a feature of the gauge symmetry fields of the SM is that they must be massless. This can be seen when considering that a Lagrangian mass term for a fermion field takes the form $m^2\bar{\psi}\psi$ which can then be decomposed into both left and right handed chiral fields as

$$m^2(\bar{\psi}_L + \bar{\psi}_R)(\psi_L + \psi_R) = m^2(\bar{\psi}_L\psi_L + \bar{\psi}_R\psi_R + \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L). \tag{2.2}$$

This expression contains terms which couple left and right-handed fermion fields together which individually are known to have different transformation properties under the weak force.

This means that the existence of these terms would break gauge symmetry and so they are not permitted by Nature. A similar problem is seen for bosons where the equivalent mass term is not invariant under a boson field gauge transformation.

For both the strong and electromagnetic interactions this presented no obvious theoretical problem due to both the photon and gluon having no mass. However, as the $W^{\pm}$ and Z^0 are both massive, the question arose of how a non-zero mass appears within a gauge field theory. This problem was solved in 1964 with the Higgs mechanism [60–63] which provides a way to break the electroweak symmetry of the universe and simultaneously give rise to gauge-compatible mass terms. In addition, it eventually leads to the introduction of quark flavour changing within the SM. Mathematically, the mechanism introduces a new term to the SM Lagrangian:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{Higgs}} \tag{2.3}$$

where, $\mathcal{L}_{\text{EW}}$ and $\mathcal{L}_{\text{QCD}}$ contain the dynamics of the electroweak and QCD aspects of the SM, while the additional term can be written as

$$\mathcal{L}_{\text{Higgs}} = (D^{\mu}\Phi)^{\dagger}(D_{\mu}\Phi) - V(\Phi), \qquad \Phi = \begin{pmatrix} \phi^{+} \\ \phi^{0} \end{pmatrix} \tag{2.4}$$

where Φ represents a new complex scalar field. The D_{μ} operators are covariant derivatives given by

$$D_{\mu} = \partial_{\mu} + igT^{i}W_{\mu}^{i} + i\frac{1}{2}g^{'}B_{\mu} \tag{2.5}$$

which contain terms for the three $SU(2)_L$ gauge bosons, W_{μ}^{i}, and the single $U(1)_Y$ gauge boson, B_{μ}. T^{i} are the Pauli matrices while both g and $g^{'}$ are coupling terms.

The potential energy of the vacuum, $V(\Phi)$, can be written in the form

$$V(\Phi) = -\mu^2\Phi^{\dagger}\Phi + \lambda\Big(\Phi^{\dagger}\Phi\Big)^2 \tag{2.6}$$

and finding the minimum of $V(\Phi)$ yields a vacuum energy expectation value (VEV), v, often parameterised in terms of the coupling constants as

$$v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{2.7}$$

For cases where $\mu^2 > 0$ there are an infinite number of potential minima lying around a circle in the complex plane of Φ as shown in Figure 2.2.

In order to break the $SU(2)_L \times U(1)_Y$ electroweak gauge symmetry, while leaving a

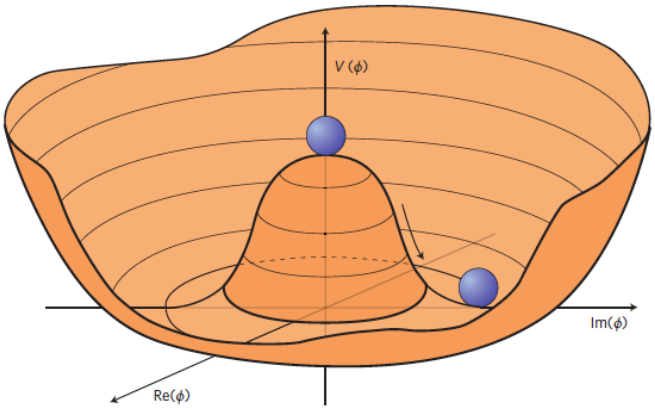

Figure 2.2: The so-called "Mexican-hat" potential of the vacuum [64].

$U(1)_Q$ electromagnetic gauge symmetry unbroken, a form of Φ which is a neutral field and therefore invariant under $U(1)_Q$ is of interest. Canonically, this form is chosen to be:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.8}$$

with oscillations about the potential minimum included by perturbing v to $v+h$ [44]. Substituting the perturbed VEV back into the kinetic term of Equation 2.4 yields terms quadratic in h, quadratic in the combination of $W^{1\mu}$ and $W^{2\mu}$ and quadratic in the combination of $W^{3\mu}$ and B^{μ}. The results, therefore, are that a new massive scalar boson, the now named Higgs boson, is born. Additionally, the $W^{1,2\mu}$ mix to provide mass terms for the $W^{\pm}$ bosons while the third gauge boson, $W^{3\mu}$, mixes with the $U(1)$ gauge boson, B^{μ}, to provide mass for the Z^0 boson. Meanwhile the electromagnetic gauge field, A^{μ}, lacks any quadratic term meaning it is left massless. A diagram of the boson mixing with the gauge fields defined as a function of the Weinberg mixing angle, θ_W, is shown in Figure 2.3.

Counting degrees of freedom before and after the electroweak symmetry breaking shows that the theory has not thrown away any physical information. Before symmetry breaking, there were twelve degrees of freedom. The SU(2) doublet is complex meaning it had four degrees of freedom. The massless $SU(2)_L \times U(1)_Y$ gauge bosons, of which there were four, each had a further two giving the final total of twelve. After the vacuum gains a VEV we have a real scalar Higgs boson with one degree of freedom, the now massive $W^{\pm}$ and Z^0 bosons each with three degrees of freedom and a massless photon with two transverse polarisations, again giving a total of twelve.

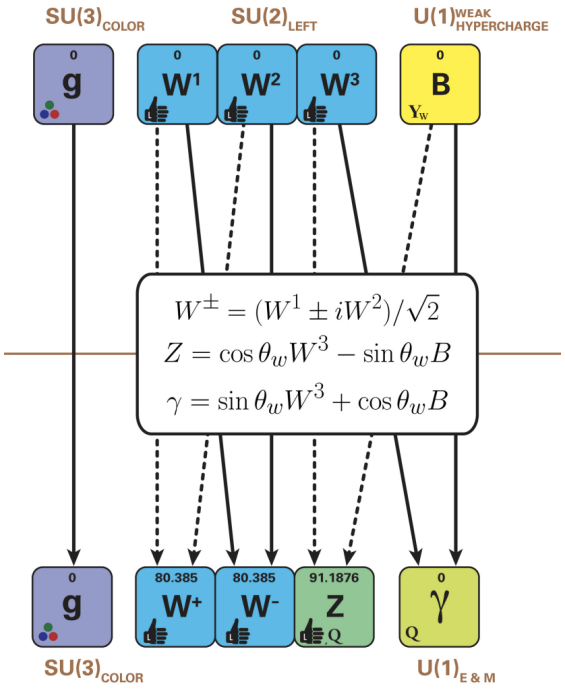

Figure 2.3: The mixing of the $SU(2)_L$ and $U(1)_Y$ gauge bosons during spontaneous electroweak symmetry breaking and the resulting gauge bosons of the SM. The mass fields are defined and the figure is edited from the original taken from [65].

2.5 Quark Masses and Flavour Changing Transitions

The Higgs mechanism which solves the issue of bosonic mass terms within the SM, also provides mass terms for the quarks. Interestingly, this source of quark masses is intrinsically linked to the ability of the quarks to change flavour. To demonstrate this firstly it is recalled that within the SM chirally left-handed quarks are arranged into SU(2) doublets, Q_{Li}, whilst chirally right-handed quarks are arranged into SU(2) singlets, u_{Rj} or d_{Rj}, i.e.

$$Q_{Li} = \begin{pmatrix} u \\ d \end{pmatrix}_L \text{ or } \begin{pmatrix} c \\ s \end{pmatrix}_L \text{ or } \begin{pmatrix} t \\ b \end{pmatrix}_L \tag{2.9}$$

and

$$\begin{aligned} u_{Rj} &= u_R \text{ or } c_R \text{ or } t_R, \\ d_{Rj} &= d_R \text{ or } s_R \text{ or } b_R. \end{aligned} \tag{2.10}$$

A new term, $\mathcal{L}_Y$, can be added to the SM Lagrangian which is of the form

$$\mathcal{L}_Y = -Y^d_{ij}\overline{Q_{Li}}\Phi d_{Rj} - Y^u_{ij}\overline{Q_{Li}}\epsilon\Phi^* u_{Rj} \tag{2.11}$$

where $Y^{u,d}$ are the 3×3 Yukawa matrices, ϵ is the anti-symmetric Levi-Civita tensor and Φ is the Higgs field. By substituting in the VEV from Equation 2.8 mass terms arise of the form

$$\frac{v}{\sqrt{2}}Y^d_{ij}\overline{q_L}q_R \tag{2.12}$$

where q denotes either an up or down type quark. The Yukawa mass matrix terms are in the flavour basis and, in this consideration of mass terms, should be transformed to the mass basis. This is done by diagonalising the Yukawa matrix terms using unitary matrices denoted by, $V_{qL,R}$:

$$\frac{v}{\sqrt{2}}Y^d \rightarrow \frac{v}{\sqrt{2}}V_{qL}Y^d V^\dagger_{qR} \tag{2.13}$$

with the same shift of basis performed for the quark terms:

$$q_{iL} \rightarrow V_{qL}q_{iL}. \tag{2.14}$$

Considering the current corresponding to the $W^\pm$ gauge bosons, $J^\mu_{W^\pm}$, and substituting in the change of bases gives:

$$J^\mu_{W^\pm} = \frac{1}{\sqrt{2}}\overline{u_{iL}}\gamma^\mu d_{iL} \rightarrow \frac{1}{\sqrt{2}}\overline{u_{iL}}\gamma^\mu(V_u^\dagger V_d)d_{iL} \tag{2.15}$$

where the matrix combination

$$(V_u^\dagger V_d) = V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.16}$$

is known as the CKM matrix, named after Cabibbo, Kobayashi and Maskawa [66]. The CKM matrix terms provide a coupling between the different quark flavours and, therefore, both the quark masses and the phenomenon of quark flavour changing is revealed simultaneously. Conversely, as a neutral current from the Z^0 boson contains terms only of the form $\overline{q^i}\gamma^\mu q^i$ there are no terms which give rise to the CKM matrix upon the change of basis, meaning that interactions involving the Z^0 boson do not change flavour in the SM.

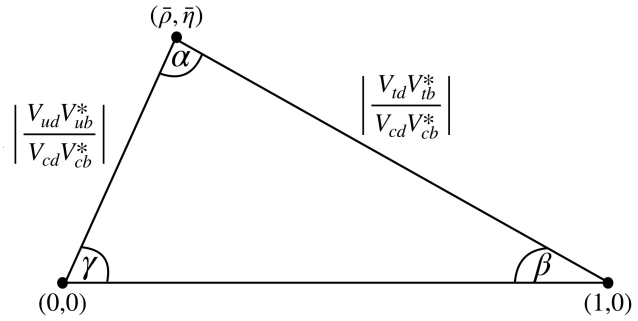

Figure 2.4: The commonly represented CKM triangle with an apex at $(\overline{\rho},\overline{\eta})$. This triangle is commonly used as the three mixing angles, position of the corners of the triangle and the length of the sides are easier to visualise.

2.6 Parameterisation of the CKM Matrix

The CKM matrix has four parameters: three real parameters and a single complex phase. This stems from the fact that any complex 3×3 matrix has 18 parameters. Half of these parameters are removed by insisting that the matrix is unitary. Another 5 of the degrees of freedom are removed by imposing phase rotations on the six flavours of quark. The four remaining parameters can be represented by three rotation (Euler) angles and a single complex phase. Crucially, it is the overall complex phase left over which is responsible for all CP violation within the SM.

A useful parameterisation, the *Wolfenstein* parameterisation [67], is commonly used as it makes explicit the matrix's hierarchy; the matrix terms further away from the diagonal which correspond to inter-generational flavour changing are more suppressed that those which only change flavour within the same generation. This parameterisation is given by

$$V_{\text{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(q-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2.17}$$

where the current best fits give values of $A \sim 0.84$, $\lambda \sim 0.23$, $\rho \sim 0.16$, $\eta \sim 0.35$ [68].

The unitarity of the CKM matrix ensures that rotations can be made between bases without changing any of the underlying physics. This feature enables the construction of several unitarity relations such as

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.18}$$

which trace out triangles in the complex plane. In total there are six possible triangles to construct from the unitarity relations however the most commonly used is as represented in Figure 2.4 The

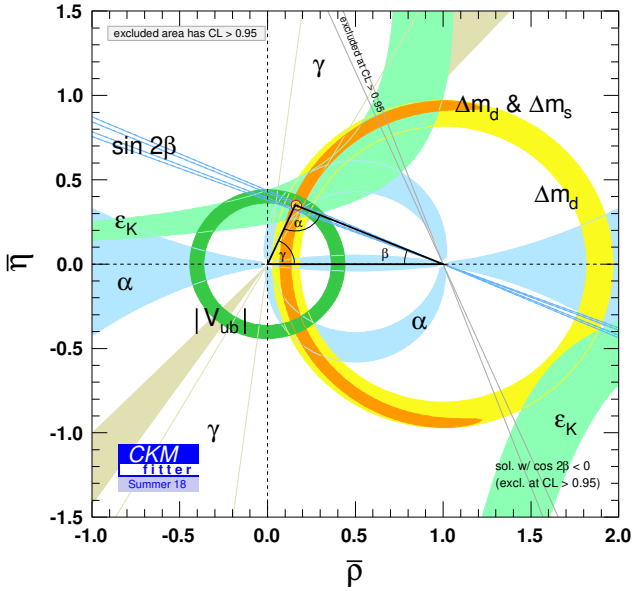

Figure 2.5: Fit to the standard CKM triangle taken from the most recent global results from CKM fitter [68]. The impressive agreement on the location of the apex can be seen from the overlapping bands.

area for all of the CKM triangles is the same and is given by half of the Jarlskog invariant [69], J, where

$$J \propto \sin(\delta), \tag{2.19}$$

where δ is the complex phase of the CKM matrix. As it is a function of δ, the CP violating parameter within the CKM matrix, the area of the CKM triangle provides a direct measure of the amount of CP violation within the SM. Measuring the angles of the CKM triangles is therefore of great interest and, over the last 15 years, there has been an enormous improvement in the precision of the measured parameters thanks to the work of both LHCb and the B-factory experiments, BABAR and Belle. The current best fit of the CKM triangle can be seen in Figure 2.5.

2.7 Flavour Changing Neutral Currents

Quark flavour is changed via charged current interactions but whether such a flavour change can take place with a neutral current interaction was a question only answered during the discovery of the need for a second generation of quarks in 1970 [70]. Faced with the lack of experimental evidence for the existence of the decay $K^0_{\mathrm{L}} \to \mu^+\mu^-$, Glashow, Iliopoulos and Maiani introduced

the charm quark to complete the second generation of quarks alongside the already known strange quark. This prediction explained the apparent absence of the $K_{\mathrm{L}}^{0} \to \mu^{+}\mu^{-}$ process by introducing the idea that, at tree level, a flavour changing neutral current (FCNC) has a contribution from the first quark doublet which cancels with that from the second, resulting in tree level FCNCs being rendered impossible. At loop level there is not a complete cancellation, but a heavy suppression with the sum of the two decays producing a small contribution of order

$$\frac{g^4(m_c^2 - m_u^2)}{m_W^2} \sim \frac{\alpha^2 m_c^2}{m_W^2} \tag{2.20}$$

which, in the case of $K_{\mathrm{L}}^{0} \to \mu^{+}\mu^{-}$, results in a small branching fraction of $\mathcal{O}(10^{-9})$ [71], explaining it's elusiveness. The existence of the FCNC terms at loop level relies, as can be seen from Equation 2.20, on non-degenerate intermediate quark masses. The GIM mechanism extends from two to three generations and, due to the masses of the quarks within the SM and the large mass of the top quark, FCNC processes are less suppressed when involving an external down-type quark than an external top-type quark.

The GIM mechanism provides an additional layer of suppression to FCNC processes in addition to the CKM matrix terms, making them very rare within the SM. Their rarity makes them a valuable tool in the hunt for new physics for several reasons. The presence of new physics within the loop would lead to a measurable alteration of certain physical observables such as the rate of the decay or variables associated to the angular structure of the decay. Additionally, as the SM predictions of the rate of FCNC processes are very small, any background from the SM would also be small. Therefore, the effect of the new physics would be comparatively large. Specifically, the small SM contributions result in an expected sensitivity to new physics up to energies of around $10\,\mathrm{TeV}/c^2$ [72]. This makes the job of hunting for a new physics signal slightly less challenging than it would otherwise be.

2.8 Characterising New Physics with Effective Field Theory

If FCNC processes are to be used to hunt for new physics it becomes necessary to remember that the SM is a low-energy effective theory. Indeed, the SM is expected to break down at high energies such as the Planck scale where the strength of the gravitational force becomes comparable to the electroweak and strong forces. Furthermore, theoretical models fail to predict the convergence of the coupling strengths of the fundamental forces to a single point in the high energy limit despite the assumption that this should happen in Nature [73]. In general, effective field theories within particle physics make use of the fact that the different particle masses are spread over many orders of magnitude. As a result, physical processes which take place at a high

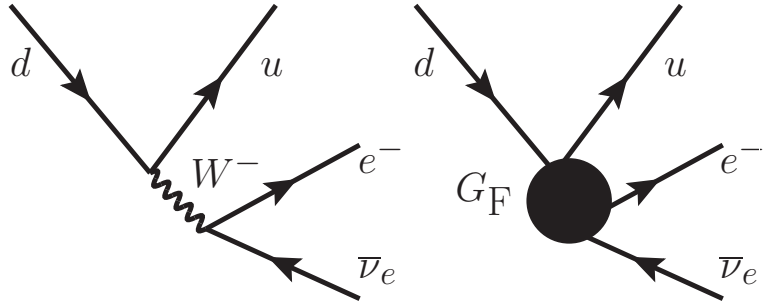

Figure 2.6: Feynman diagram for Beta decay (left) with the Fermi theory equivalent (right) which assumes that the weak interaction is a point-like four fermion interaction with a coupling strength of G_F.

energy scale are effectively decoupled from processes which take place at a lower energy scale and therefore the two processes can be treated, with care, separately [74]. Crucially, an effective field theory is only considered a valid approximation when performing calculations at a particular energy scale appropriate for the particles in question.

Arguably the most famous example of an effective field theory within particle physics is that of Fermi's theory of beta decay, proposed in 1933 [75, 76]. The theory postulated that beta decay could be well described by a four point fermion interaction, as seen in Figure 2.6, with a coupling given by, G_{F}, the Fermi constant. The theory describes weak decays very well up to energies of around 100 GeV beyond which the energy scale approaches the mass of the $W^{\pm}$ and so the accuracy breaks down. It is now understood that beta decay includes the propagator W^- term and so Fermi theory is an oversimplification. However, the predictive accuracy of the theory at lower energies provided the precedent of using effective theories elsewhere in the absence of a more fundamental theory.

When considering rare, FCNC decays involving a $b \to q$ transition an effective field theory can be used where particles with masses much larger than m_b, such as the top quark and the $W^{\pm}$ bosons, are integrated out. It is then possible to use an *operator product expansion* and define an effective Hamiltonian given by,

$$\mathcal{H}_{\text{eff}} = \frac{-4G_F}{\sqrt{2}} V_{tb} V_{tq}^* \sum_i C_i(\mu) \mathcal{O}_i(\mu) \tag{2.21}$$

which is discussed in more detail in [77, 78]. This Hamiltonian only contains CKM terms involving a top quark transition as it assumes that these terms dominate over those with intermediate up and charm quarks, due to the very large top mass and the resulting weak GIM suppression.

Analogously to Fermi theory the short distance, or high energy, physics is encapsulated in

a series of couplings, known as Wilson coefficients, $C_i(\mu)$, which are calculated at the electroweak scale by comparing the full theory to the effective theory before being evolved down to the energy scale under study [79]. In addition, the structure of the fermion interaction in question is described by a series of matrix operators, $\mathcal{O}_i(\mu)$. Of particular interest are the operators $\mathcal{O}_{7,9,10}$, which provide contributions to semi-leptonic $b \to qll$ processes. They are defined as:

$$\mathcal{O}_7 = \frac{e}{16\pi^2} m_b (\overline{q}_L \sigma_{\mu\nu} b_R) F^{\mu\nu} \tag{2.22}$$

$$\mathcal{O}_9 = \frac{e^2}{16\pi^2} m_b (\overline{q}_L \gamma_\mu b_L) \sum_l (\bar{l}\gamma^\mu l) \tag{2.23}$$

$$\mathcal{O}_{10} = \frac{e^2}{16\pi^2} m_b (\overline{q}_L \gamma_\mu b_L) \sum_l (\bar{l}\gamma^\mu \gamma_5 l) \tag{2.24}$$

where $\mathcal{O}_7$ describes radiative decays, and $\mathcal{O}_{9,10}$ describe decays via a semi-leptonic vector and axial vector current respectively. Additionally, for each operator, one can also define a primed set of operators, $\mathcal{O}'$, which correspond to the chirally-flipped state. Due to the left handedness of the weak interaction the corresponding Wilson coefficients for these operators, C_i', are suppressed by a factor of m_q/m_b. Despite this suppression, $\mathcal{O}'_{7-10}$ could be greatly enhanced by the presence of new physics [80].

It is possible that there could be new contributions with the inclusion of new, scalar, pseudo-scalar or tensor operators. There are many possibilities of new operators such as $\mathcal{O}_{S,P}$ both of which are zero in the SM but could arise in new physics scenarios. Indeed, the consequences of including new operators which describe beyond the SM processes could be large [72]. For example, the addition of a new operator which is complex would point to the discovery of a new source of CP violation in Nature. In addition, it is conceivable that new operators could exist which result in lepton flavour violation, or which do not respect the lepton flavour universality which is present in the SM couplings between gauge bosons and leptons.

Ultimately, searching for new physics via a particular Wilson coefficient is equivalent to testing a specific form of Lorentz structure as the Wilson coefficient is independent of any specific process. As a result a statistically significant deviation from the SM seen in global fits to Wilson coefficients informs which class of process should be studied further. Current Wilson coefficient fits are discussed in more detail in Section 2.10.

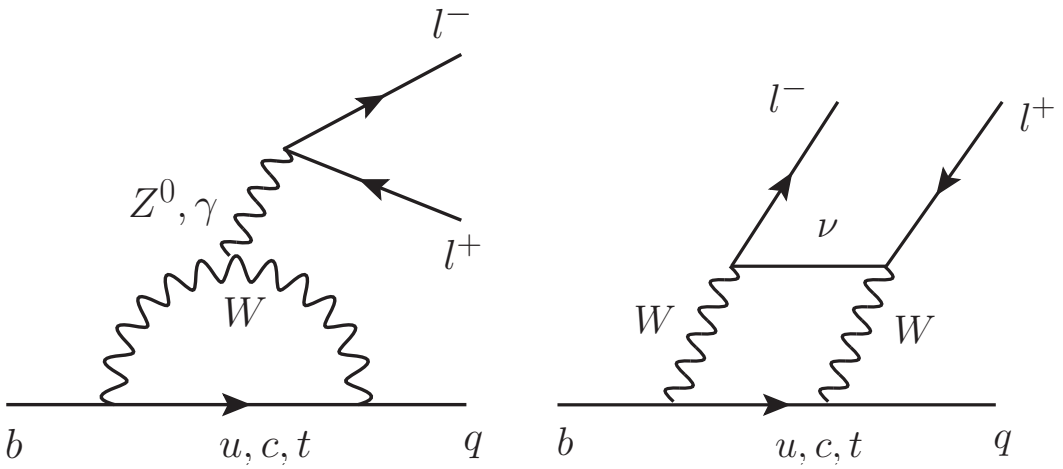

Figure 2.7: Lowest level FCNC Feynman diagrams for a $b \to q l^+ l^-$ process.

2.9 $b \to q l^+ l^-$ processes

As mentioned previously FCNC decays with external top-type quarks suffer from heavier GIM suppression than those with external down-type quarks. Indeed, top-type FCNC processes which take place via a $t \to cg$, $t \to c\gamma$ or $t \to cZ$ process have predicted branching fractions of $\mathcal{O}(10^{-10})$, $\mathcal{O}(10^{-12})$ and $\mathcal{O}(10^{-13})$ respectively [81], making measurements of these modes experimentally difficult. The FCNC processes of the form $b \to q l^+ l^-$ have larger predicted branching fractions, of $\mathcal{O}(10^{-6} - 10^{-9})$ (e.g. [82]) and, when coupled with the relatively long lifetime of B mesons at $\mathcal{O}(10^{-12})$ seconds, makes them a more promising candidate for experimentation. Examples of lowest level diagrams for these processes are shown in Figure 2.7.

A feature of these processes is that the contributions of the operators $\mathcal{O}_{7,9,10}$ to the rate of the decay are known to vary depending on the decay kinematics. Specifically, the variation of the decay rate with the value of q^2, which is defined as the invariant mass squared of the dilepton system [83], is of particular interest, as shown in Figure 2.8. Probing a specific region of q^2 leads to an increased sensitivity to particular Wilson coefficients and, by extension, specific classes of underlying physics. Figure 2.8 also shows large peaks which originate from the narrow charmonium resonances, J/ψ and $\psi(2S)$. These point to long distance $b \to q c \bar{c}$ contributions that decay to two charged leptons and which obscure any sensitivity to the potential new physics. As a result, such regions are normally vetoed in any $b \to q l^+ l^-$ study.

Another motivation for performing $b \to q l^+ l^-$ analyses in particular regions of q^2 arises from the handling of uncertainties associated to the calculation of QCD hadronic terms. The operator term can be decomposed and factorised into two components representing the leptonic

and hadronic parts of the process:

$$\begin{aligned}\langle X l\bar{l}|\mathcal{O}|B^0\rangle &= \langle X|I_{\text{Quark}}|B^0\rangle \cdot \langle \bar{l}l|I_{\text{lepton}}|0\rangle + \text{non-factorisable terms} \\ &= F(q^2) \cdot \langle \bar{l}l|I_{\text{lepton}}|0\rangle + \text{non-factorisable terms}\end{aligned} \tag{2.25}$$

where X denotes a hadron, the I terms the quark and lepton currents, $F(q^2)$ a hadronic form-factor and where non-factorisable terms refer to cases where the assumption of decoupled short and long distance physics breaks down. The calculation of the hadronic form-factor as a function of q^2 provides the dominant source of systematic uncertainty in the theoretical predictions for these decays due to the non-pertubative nature of QCD at low energies. Typically, these large uncertainties are around 30% of the central value [84–86]. At high hadronic recoil, corresponding to a low q^2, sum rules which exploit the analyticity of the equations on the light cone [87] are used to calculate these form-factors (e.g. [88]). At low recoil, where the QCD coupling strength is highly non-linear, QCD calculations can be performed using a space-time lattice with some finite element spacing, a [89, 90]. As $a \to 0$ and the size of the lattice itself $\to \infty$ the QCD continuum becomes well described up to a momentum cut-off of the order of $\frac{1}{a}$ meaning QCD

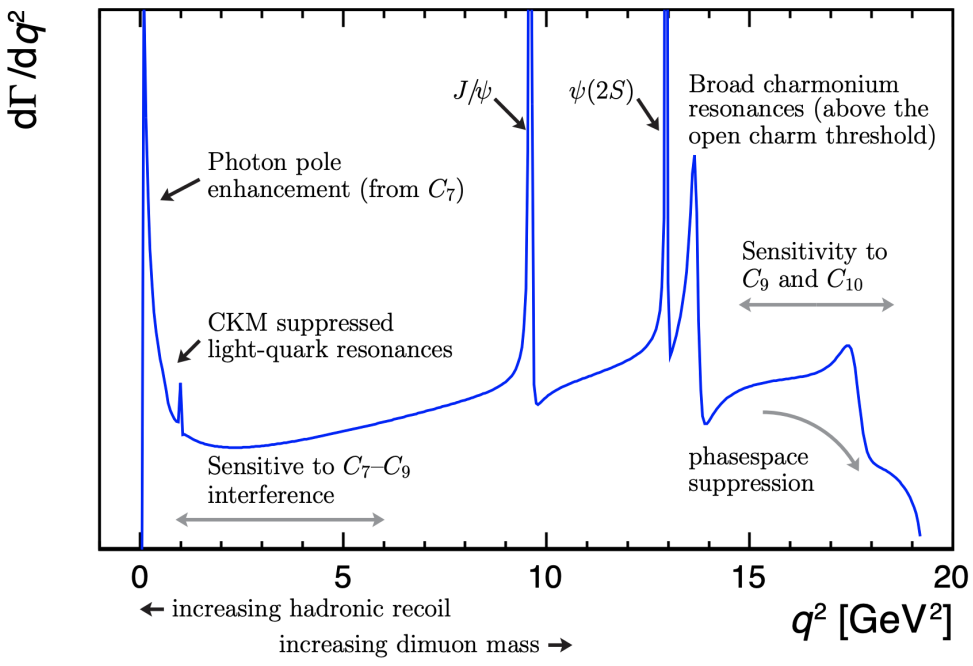

Figure 2.8: Differential decay rate of a $b \to q l^+ l^-$ process as a function of the invariant mass of the dilepton pair, q^2. The regions dominated by particular Wilson coefficient contributions are labelled. Figure taken from Ref. [83]

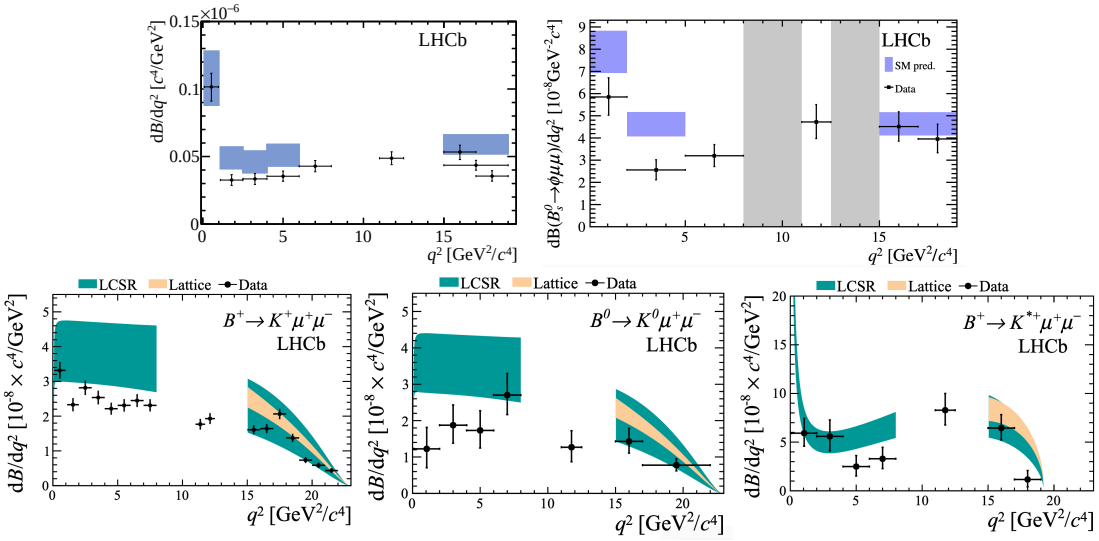

Figure 2.9: Differential branching fraction of $B^0 \to K^{*0}\mu^+\mu^-$ [91], $B^0_s \to \phi\mu^+\mu^-$ [92], $B^0 \to K^+\mu^+\mu^-$, $B^0 \to K^0\mu^+\mu^-$ and $B^0 \to K^{*+}\mu^+\mu^-$ [93].

calculations can be performed.

2.10 Experimental status of $b \to sl^+l^-$ processes

There are, to date, some interesting tensions with the SM arising in the study of $b \to sl^+l^-$ processes as measured by both LHCb and the B-factory experiments, Belle and BABAR. These tensions are present in several complementary physical measurements, and, most excitingly, seem to occur within the same low q^2 region.

The first of these tensions has arisen in the measurements of differential branching fractions of several decay modes including $B^0_s \to \phi\mu^+\mu^-$, $B^0 \to K^*\mu^+\mu^-$, $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^0\mu^+\mu^-$ and $B^+ \to K^{*+}\mu^+\mu^-$ [91–93]. Despite the rarity of these processes, (for example $B^0 \to K^{*0}\mu^+\mu^-$ has a branching fraction of $\mathcal{O}(10^{-7})$), the large amount of data taken by the LHCb experiment has made measurements in bins of q^2 experimentally feasible. Both the latest experimental results alongside the theoretical predictions are shown in Figure 2.9.

The set of results show, consistently, that the SM prediction for the branching fraction at low values of q^2 seems too high when compared to the experimental results. Although in each decay mode this discrepancy does not reach a 5 σ confirmation threshold, the fact that the same effect is seen over multiple decay modes is, perhaps, a clue of some underlying pattern.

Unfortunately, the SM predictions for the differential branching fractions suffer from the

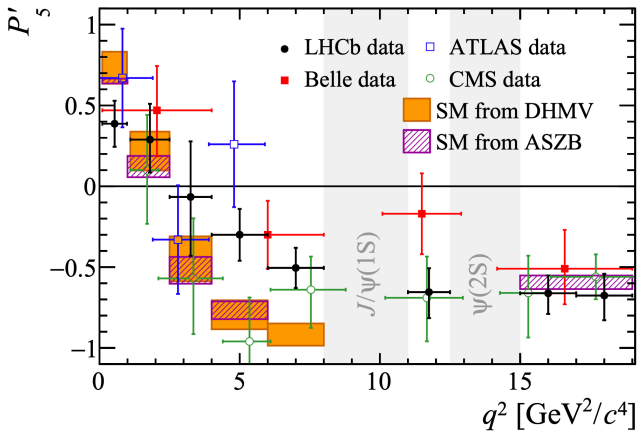

Figure 2.10: The P'_5 LHCb measurements [95], ATLAS [96], CMS [97], Belle [98], and theory predictions from [99] and [100]

large hadronic uncertainties mentioned previously. However, a second source of tension with the SM, which stems from the angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$, is sensitive to angular observables which have reduced form-factor uncertainties and are thus a cleaner probe for new physics [94]. The angular analysis describes the decay in terms of the helicity angles of the K^* and dilepton systems and the angles between decay planes. It is possible to parameterise different physical observables in terms of the transversity amplitudes of the K^* meson and a subset of these observables have been "optimised" to reduce form-factor uncertainties. This is achieved by constructing the optimised observables as ratios of transversity amplitudes, and hence ratios of form-factors. This exploits the fact that by expanding the form-factor ratios to first order in $1/m_b$ the corresponding form-factor uncertainties will cancel as many of the form-factors are related at low q^2. An example of such an observable is P'_5 which is defined as

$$P'_5 = \sqrt{2}\frac{\mathcal{R}(A_0^L A_\perp^{L*} - A_0^R A_\perp^{R*})}{\sqrt{(|A_0^L|^2 + |A_0^R|^2)(|A_\parallel^L|^2 + |A_\perp^L|^2) + |A_\parallel^R|^2 + |A_\perp^R|^2)}} \tag{2.26}$$

where A_0 refers to a longitudinally polarised amplitude and $A_\parallel$ and $A_\perp$ to transversely polarised amplitudes. The measurement of P'_5 in bins of q^2 from LHCb, ATLAS and CMS with Run 1 LHC data and from Belle can be seen in Figure 2.10. Once again, there is a tension with the SM theory within the low q^2 region, specifically a 2.8 σ tension in the $q^2 \in [4.0, 6.0]$ GeV$^2/c^4$ region and a 3.0 σ tension in the $q^2 \in [6.0, 8.0]$ GeV$^2/c^4$ region.

The tensions in both the differential branching fraction and P'_5 result could potentially be explained by hadronic effects not accounted for within the theoretical estimates such as the inclusion of interactions with additional charm loops outside of the charmonium regions [101]. Addi-

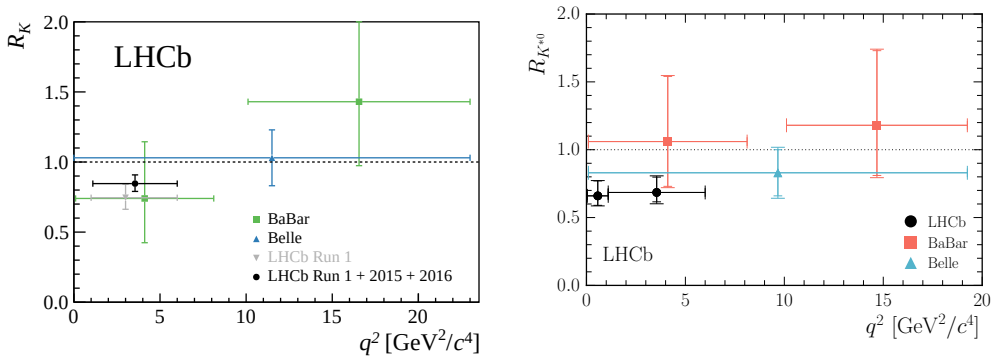

Figure 2.11: Measurements of lepton universality test ratios, R_K (left) and $R_{K^{*0}}$ (right). LHCb measurements for R_K are provided using Run 1 [105] and 2015, 2016 data [104], alongside Run 1 data for $R_{K^{*0}}$ [106]. R_K and $R_{K^{*0}}$ measurements are also provided by Belle [108] and BABAR [107].

tionally, it is possible that the current theoretical uncertainties associated to the non-factorisable hadronic terms where, for example, a spectator quark interacts with the short distance physics, may be underestimated across the q^2 range [102, 103]. Understanding these hadronic effects is therefore a prerequisite to being able to claim a new physics discovery with these measurements alone. This theoretical barrier is not present, however, for a final set of tensions which are seen in studies that test the universality of the electroweak coupling strength for the different flavours of charged lepton. Within the SM it is expected that the different lepton flavours couple with the same strength and, therefore, that the ratio

$$R_X = \frac{\Gamma(B^0 \to X\mu^+\mu^-)}{\Gamma(B^0 \to Xe^+e^-)} \tag{2.27}$$

is equal to unity (differing only due to mass differences between the leptons). The hadronic form-factor uncertainties that prove problematic for the branching fraction and, to a lesser extent, the angular analyses are here effectively removed by the ratio. Other systematic effects will also cancel in the ratio meaning that these tests of the SM are very clean from a theoretical standpoint. As shown in Figure 2.11, the processes with $X = K^+, K^{*0}$ show a 2.5 σ and 2.1-2.5 σ (differs depending on theoretical models) tension with the SM respectively within the low q^2 range [104–108]. Furthermore, these results complement the so-called $R_{D^{(*)}}$ measurements made by the B-factories and LHCb which, while considering $b \to cl\nu_l$ processes instead of $b \to sll$ processes, also show tensions with the SM expectation at a combined significance of 3.8 σ [109–115].

Ultimately, additional data from LHCb and from the future Belle 2 experiment is required to determine whether these tensions are true smoking gun signals of new physics. With the current

datasets however, global fits in WC space can be made by using the operator product expansion to convert results into WCs and then performing a global likelihood fit using results from the different measurements. At the moment these global fits show a preference for the inclusion of new physics with modifications to C_9, and to a lesser extent, C_{10} preferred. The latest C_9 and C_{10} fits show a 4 σ tension with the SM while the quality of the fits is seen to improve with the addition of small lepton universality new physics [116]. Examples of such global fits are shown in Figure 2.12. Some of the new physics models which are currently proposed to address the B decay anomalies include the introduction of a new, Z' vector boson (e.g. [117]), or models which involve a U(1) leptoquark (e.g. [118, 119]). Both models suggest that the inclusion of a scalar leptoquark/Z' can simultaneously explain the lepton non-universality results as the models have different lepton family couplings built in. Alongside the B anomalies the models also claim a possible explanation for other tensions seen in particle physics, such as the famous anomalous magnetic moment of the muon, $(g-2)_\mu$ [120].

2.11 $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ as examples of $b \to d(s)l^+l^-$

If new physics is indeed discovered within the $b \to s$ processes then its flavour structure could be probed by studying the similar and yet, to date, less studied $b \to dl^+l^-$ processes. These experience even greater CKM suppression due to the small size of the off-diagonal CKM matrix element $|V_{td}|$ meaning that $b \to dl^+l^-$ processes are potentially even more sensitive to new physics signals. Furthermore, the study of such modes would provide a new source of constraint on the CKM triangle with measurements of $|V_{td}|/|V_{ts}|$ able to be extracted from data.

The $b \to dl^+l^-$ processes are also potentially sensitive to large CP violating effects. This is due to the fact that the relevant CKM amplitudes for a $b \to d$ process contain the contributions $V_{tb}V_{td}^*$, $V_{cb}V_{cd}^*$, $V_{ub}V_{ud}^*$ which are all proportional to λ^3 where λ is the Wolfenstein CKM parameter. As the contributions are all of the same magnitude and each have different weak phases, the interference effects between the different processes could be large. The equivalent terms for $b \to s$ processes differ from one another by orders of magnitude in λ and therefore are not as interesting for these CP violation searches.

The primary focus of this thesis is the decay $B^0 \to \pi^+\pi^-\mu^+\mu^-$, which is an example of a $b \to d\mu^+\mu^-$ process believed to mainly proceed via the decay $B^0 \to \rho^0(770)\mu^+\mu^-$ where $\rho^0(770) \to \pi^+\pi^-$ [121] and where the $\rho^0(770)$ is a vector meson comprising of a superposition of up and down quarks. Theoretical predictions for the branching fraction are of $\mathcal{O}(10^{-8})$. Predictions for the full q^2 range include 2.4×10^{-8} and 4.1×10^{-8} where these different values originate from the different form factors used in the theoretical calculations [84–86]. These predictions suffer from the large hadronic form-factor uncertainties previously discussed with

uncertainties of around 30%. Theoretical predictions in the low q^2 region include $(4.2 \pm 0.5) \times 10^{-9}$ [86] while predictions in the upper q^2 region are yet to be determined.

A previous analysis by the LHCb collaboration, using the Run 1 dataset, measured a branching fraction value of

$$\mathcal{B}(B^0 \to \pi^+\pi^-\mu^+\mu^-) = (2.11 \pm 0.51\,(\text{stat}) \pm 0.15\,(\text{syst}) \pm 0.16\,(\text{norm})) \times 10^{-8} \quad (2.28)$$

at a statistical significance of 4.8 σ. The analysis was performed by selecting candidates where the invariant $\pi^+\pi^-$ mass was close to the $\rho^0(770)$ resonance [122]. The work presented within this thesis provides an updated branching fraction measurement using both Run 1 and Run 2 LHCb data.

The previous Run 1 LHCb analysis also provided the first observation of the $b \to s\mu^+\mu^-$ mode, $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ where the branching fraction was measured to be

$$\mathcal{B}(B_s^0 \to \pi^+\pi^-\mu^+\mu^-) = (8.6 \pm 1.5\,(\text{stat}) \pm 0.7\,(\text{syst}) \pm 0.7\,(\text{norm})) \times 10^{-8} \quad (2.29)$$

at a statistical significance of 7.2 σ.

The decay $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ proceeds mainly via $B_s^0 \to f_0(980)\mu^+\mu^-$. Theoretical predictions for the branching fraction across the full q^2 range include $(5.21^{+3.23}_{-2.06}) \times 10^{-7}$ [123], $(9.5^{+3.1}_{-2.6}) \times 10^{-8}$ [124] and between $(0.81 - 2.02) \times 10^{-8}$ [125] where the different predictions arise from different parameterisations and form factor values. Theoretical predictions across low and high q^2 lie in the ranges $(0.49 - 1.29) \times 10^{-8}$ and $(0.63 - 1.57) \times 10^{-8}$ respectively [125]. The already large form-factor uncertainties on these predictions are made larger by additional uncertainties due to the limited understanding of the $f_0(980)$ quark content. Although it is generally assumed that the $f_0(980)$ is a meson with a superposition of up, down and strange quarks, there is speculation that the $f_0(980)$ could in fact be a tetraquark due to it's unusual quantum numbers [126]. Being able to minimise both the statistical and systematic uncertainties on the branching fraction measurement is therefore of great interest and so this thesis also provides an updated measurement of the B_s^0 mode with the inclusion of the Run 2 LHCb data.

Given the wider context, this thesis presents the branching fraction measurements in bins of q^2, allowing for comparisons to the discrepancies seen in the $b \to s$ sector. It is hoped that by performing the measurements in specific regions of q^2, similar tensions may start to be uncovered within the $b \to d$ sector.

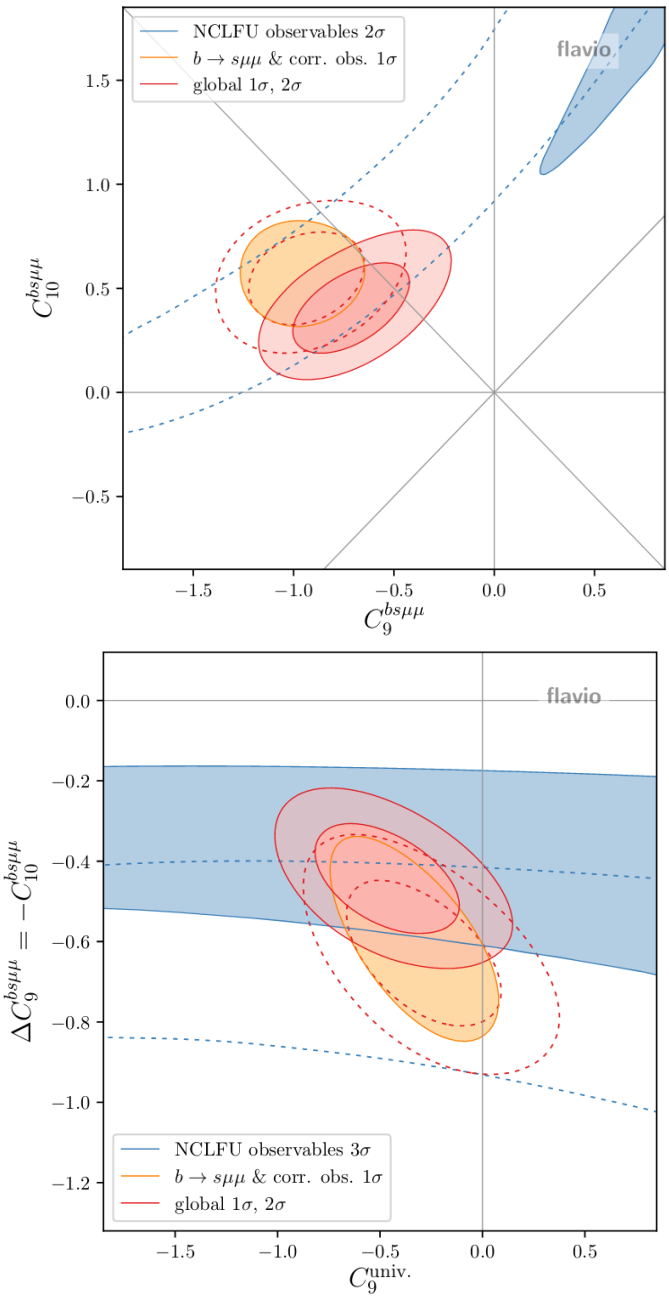

Figure 2.12: Global fits taken from Ref. [116] showing tensions seen between the SM prediction (marked by the origin) and the $b \to s\mu^+\mu^-$ results in C_9 and C_{10} (top). A sizeable preference for a modification in C_9 can be seen. Also shows (bottom) a global fit when including small lepton flavor universality new physics contributions which results in a better quality of fit. NCLFU refers to neutral current lepton flavour universality and these observables include the R_K and $R_{K^{*0}}$ measurements in addition to the P_5' measurements from the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-(e^+e^-)$.

CHAPTER 3

The LHCb detector at CERN

“ Measure what is measurable, and make measurable what is not so ”

Galileo Galilei,

This chapter provides an introduction to the LHCb experiment, which is the source of the data used within the analysis. The role of the various sub-detectors that comprise the LHCb experiment are discussed and aspects of the performance of the experiment are detailed. Details regarding the LHCb detector are taken from references [127] and [128]. The chapter begins with an introduction to the Large Hadron Collider which is the source of the proton beams that provide the particle collisions necessary for data-taking [129].

3.1 The Large Hadron Collider

Located at CERN (the European Organisation for Nuclear Research), and lying underneath the French-Swiss border near the city of Geneva, the Large Hadron Collider (LHC) holds the title of the most energetic particle accelerator in the world, and, perhaps even more impressively, also lays claim to being the largest machine that humans have ever built. The size of the LHC, at 26.6 km in circumference, is appropriate given the LHC’s fundamental objective: to attempt to help answer the fundamental questions in particle physics which have, to date, eluded understanding. In order to try and make some progress towards this ambition the LHC has been carefully designed to fulfill it’s primary functional purpose which is to provide proton-proton collisions at centre of mass energies of up to 14 TeV by accelerating two 7 TeV beams of protons around the ring

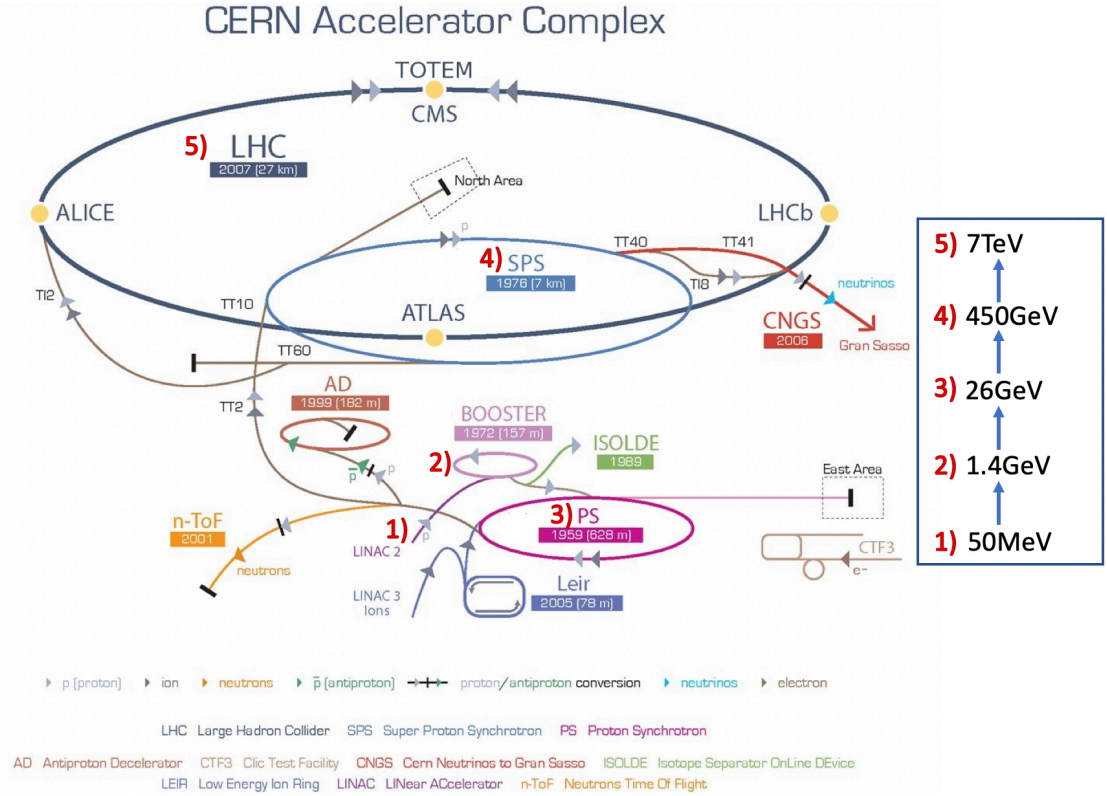

Figure 3.1: CERN accelerator complex with LINAC 2, PS Booster, PS, SPS and LHC ring labelled as 1-5. The energy of the protons after each acceleration stage is labelled. Figure is edited from CERN-DI-0606052 [130].

clockwise and counterclockwise before allowing them to collide at four different crossing points, which house different detectors.

The LHC itself is the final stage in a chain of accelerators, known collectively as the CERN accelerator complex, with each stage of the complex accelerating the particles to higher energies before injecting them into the next stage. In addition to proton-proton collisions the LHC is also designed to provide proton-ion and ion-ion collisions. The following discussion concerns the proton-proton collisions which are of interest for this thesis. The proton-proton collisions originate from a small canister of hydrogen gas whereby the gas is exposed to an electric field that strips away the electrons from the atoms. The protons are then injected into the linear accelerator, LINAC 2, where they are accelerated to an energy of 50 MeV. The protons continue, being injected into the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) being accelerated each time to 1.4 GeV, 26 GeV and

450 GeV respectively before being deposited into the LHC for a final acceleration up to 7 TeV. A schematic of the complex with acceleration energies labelled is shown for reference in Figure 3.1.

Within the LHC, charged particles are accelerated with the use of radio frequency cavities operating at a frequency of 400MHz. The particle trajectories around the ring are bent with the use of 1232 superconducting dipole magnets, each of which weighs around 35 metric tonnes and produces a magnetic field strength in excess of 8T. The magnetic field from the dipole magnets is continuously adjusted to ensure that the particles stay on their fixed orbit while the frequency of the electromagnetic field inside the cavity is synchronised to match the changing angular frequency of the protons. As a result the acceleration of the protons varies depending upon when the proton enters the cavity resulting in the formation of proton *bunches* instead of a continuous stream. The bunch structure varies between the different accelerators however, for the LHC, each bunch contains $\mathcal{O}(10^{11})$ protons with the number of bunches per beam varying between the different data taking runs, 2011-2012 (Run 1) and 2015-2018 (Run 2). The bunches are separated from each other with a minimum separation in time of 25ns which corresponds to a bunch crossing frequency of 40MHz - the operational frequency of the LHC. In addition to the dipole magnets, the LHC also makes use of quadrupole, sextupole, octupole and decapole magnets. These both compress and focus the dimensions of the beam which is necessary for beam stability and for allowing collisions at the crossing points.

The overall structure of the LHC is described in terms of sixteen separate sections. Eight are straight *insertion* sections of which four house the main LHC experiments while the other four holde either radio frequency cavities, beam dump areas or equipment used to collimate the beams. The remaining eight sections of the LHC, the *arced* sections, consist of the dipole bending magnets. In order to ensure clean particle collisions, the beam pipe is evacuated using cryogenic pumping techniques to a pressure of 10^{-7} Pa which is reduced further to 10^{-9} Pa around the beam crossing points.

A key metric to gauge the performance of any particle accelerator is the detector's instantaneous luminosity, $\mathcal{L}$ [131]. This quantity is defined as

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt} = \frac{N_B(N_p^2)f}{4\pi\delta^2}S \tag{3.1}$$

where N_B is the number of bunches in a beam, N_p is the number of protons in each bunch, δ is the transverse size of the beam at the interaction crossing point, f is the revolution frequency, S is a geometrical factor related to the beam crossing angle, σ is the cross section of the event under study and, finally, $\frac{dN}{dt}$ is the number of events generated per second. The maximum design luminosity of the LHC is $1 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ [132]. The four main experiments of the LHC, LHCb,

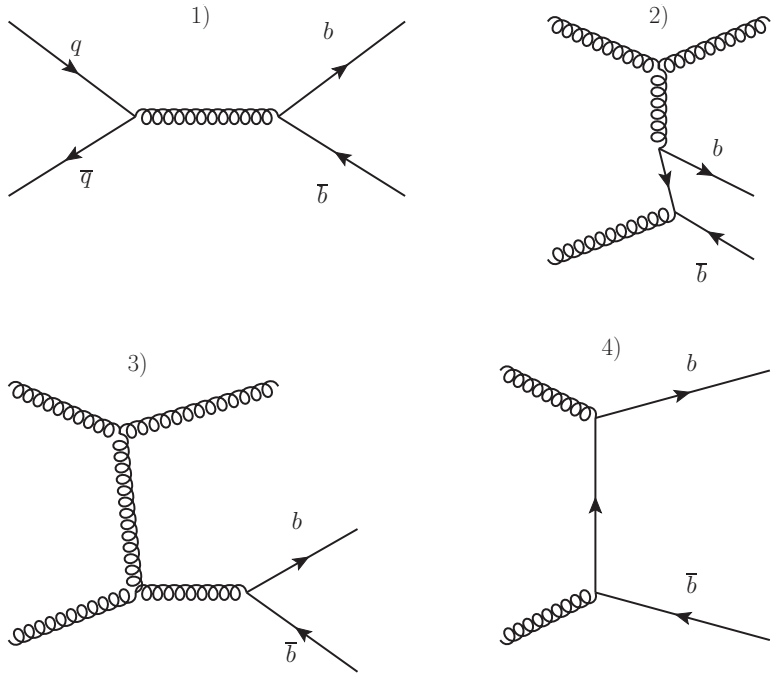

Figure 3.2: Feynman diagrams showing different channels for $b\bar{b}$ production. The mechanisms are: 1) quark-antiquark annihilation, 2) flavour excitation, 3) gluon splitting and 4) gluon-gluon fusion.

ATLAS, ALICE and CMS are each designed to run at different luminosities depending on their physics objectives. Both ATLAS (A Toroidal LHC ApparatuS) [133] and CMS (Compact Muon Solenoid) [134] are general purpose detectors which are able to study a wide breadth of high transverse momentum, (p_{T}), physics at the maximum luminosity of the LHC. ALICE (A Large Ion Collider Experiment) [135] and LHCb are designed for more specific studies with ALICE studying the quark-gluon plasmas originating from the collisions of heavy ions and LHCb CP violation and rare decay phenomena in the beauty and charm sectors. Both of these detectors are designed to operate at a lower instantaneous luminosity with ALICE's peak luminosity $2 \times 10^{29}\mathrm{cm}^{-2}s^{-1}$ and LHCb's equivalent, $2 \times 10^{32}\mathrm{cm}^{-2}s^{-1}$.

Conveniently for heavy flavour physics, the LHC produces large numbers of both $b\bar{b}$ and $c\bar{c}$ pairs from proton-proton collisions. These quark pairs are mainly produced from processes that are initiated via a gluon interaction as opposed to via valence quark scattering [136, 137] and the four main types of production are shown in Figure 3.2. Of the four mechanisms presented within the figure the prevalence of each in describing the production of $b\bar{b}$ and $c\bar{c}$ pairs varies with energy and their contributions are calculated in terms of the QCD strong coupling constant, α_s. Naively, it would be expected that the leading order processes in α_s such as the gluon-gluon fusion and

quark-antiquark processes would provide the largest contributions. However, the contributions depend on the energy of the interaction and the high energy of the LHC collisions results in the next to leading order contributions providing the largest contributions. This results in the flavour excitation processes followed by the gluon splitting processes ultimately dominating the total $b\bar{b}$ production cross-section. These are represented by sub-figures 2) and 3) of Figure 3.2.

In addition to the heavy mesons, the LHC also provides a unique laboratory to study relatively unknown b baryon decays. The fragmentation fraction of a b quark hadronising into a particular species of b hadron, X_b^i, is given by:

$$f_{X_b^i} = \frac{\sigma(b \rightarrow X_b^i)}{\sum_i \sigma(b \rightarrow X_b^i)} \tag{3.2}$$

where $f_{X_b^i}$ is the fragmentation fraction and $\sigma(b \rightarrow X_b^i)$ is the hadronisation rate. Recognising that, for example, the fragmentation fraction for the Λ_b^0 baryon is as high as approximately 40% [138] means that the LHC also provides significant scope to study baryon decays.

3.2 Introduction to the LHCb experiment

The LHCb detector is a single-arm forward spectrometer designed to probe the decay phenomena of heavy b and c quarks, which are created from the proton-proton collisions of the LHC. A general schematic of the detector, which was used for data taking between 2011-2018, can be seen in Figure 3.3. The schematic shows that the spatial extent of the detector is described using a coordinate system whereby the z-axis is aligned with the beam, the y-axis describes the vertical direction with respect to the beam and the x-axis the horizontal direction towards the centre of the LHC ring. Unlike the general purpose detectors ATLAS and CMS, LHCb has a smaller angular acceptance corresponding to the pseudorapidity range $2 < \eta < 5$, where pseudorapidity is defined as

$$\eta = -\ln(\tan(\theta/2)) \tag{3.3}$$

and where θ is the angle between a particle and the beam. This is equivalent to a 15 mrad to 300 mrad acceptance in the x direction and 15mrad to 250mrad acceptance in the y direction where the difference between the x and y acceptance takes into account the bending of the track trajectory by the LHCb magnet. The acceptance region is specifically chosen to exploit the fact that at the high energies of the LHC the production of b quarks primarily originates from the interaction of a soft parton with a hard parton meaning that the resultant $b\bar{b}$ pair is highly boosted and produced within a forward cone as shown in Figure 3.4. In total, approximately 25% of the LHC $b\bar{b}$ pairs which are produced fall within the angular acceptance of the LHCb detector.

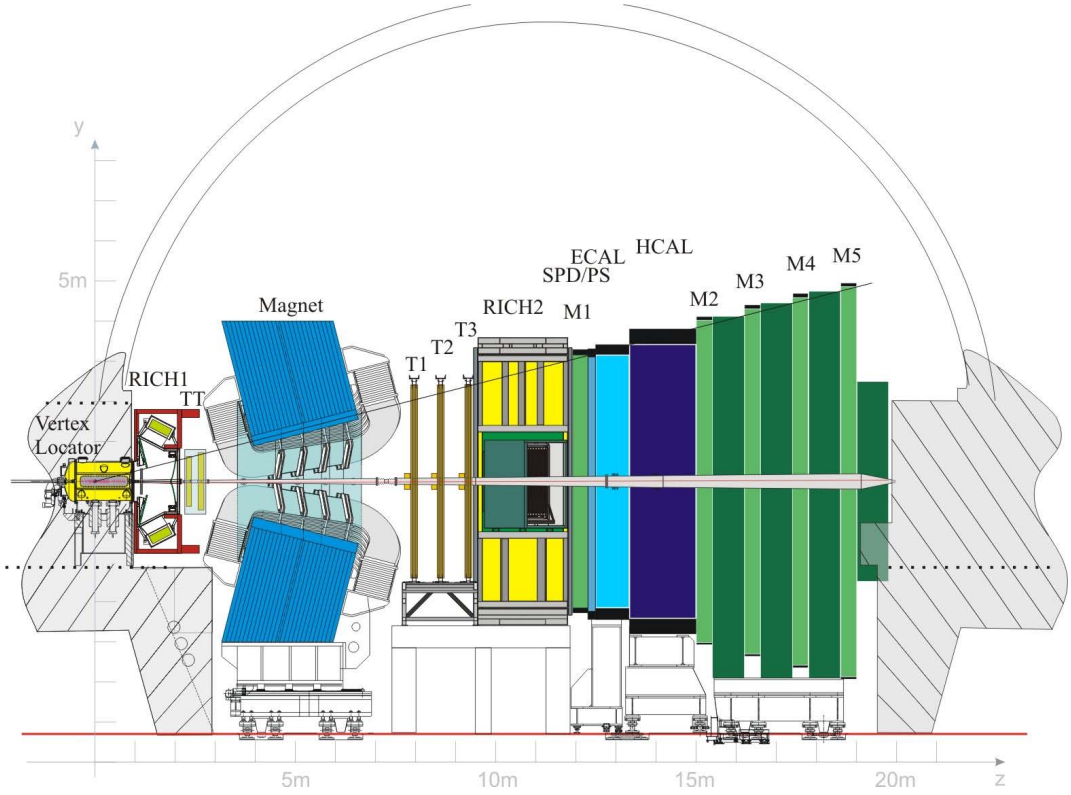

Figure 3.3: Schematic of the LHCb detector from Run 1 and Run 2 of the LHC [127] with the different sub-detectors labelled. Proton-proton collisions occur at z $= 0$.

Furthermore, the heavily boosted nature of $b\bar{b}$ pairs at the LHC means that the beauty hadrons fly a distance large enough to provide a good decay time resolution which is important for signal identification in physics analyses.

As mentioned previously the LHCb detector operates at an instantaneous luminosity of $\mathcal{O}(10^{32})\text{cm}^2\text{s}^{-1}$ that is much lower than the maximum instantaneous luminosity of the LHC. The main justification for running at this lower luminosity is that LHCb's reconstruction and trigger algorithms exploit the fact that b hadrons fly, approximately, 1 cm before decaying which is a sufficiently large distance to be resolved with the LHCb VELO sub-detector. Therefore, when reconstructing the tracks of daughter particles, a *secondary vertex* (SV) can be determined in addition to the *primary vertex* (PV) and the existence of this displaced SV is a key indicator of a heavy flavour decay. However, the process of identifying and separating the PVs and SVs becomes both much harder and also computationally intensive when there is too high a luminosity and the resulting multiple PVs per bunch crossing.

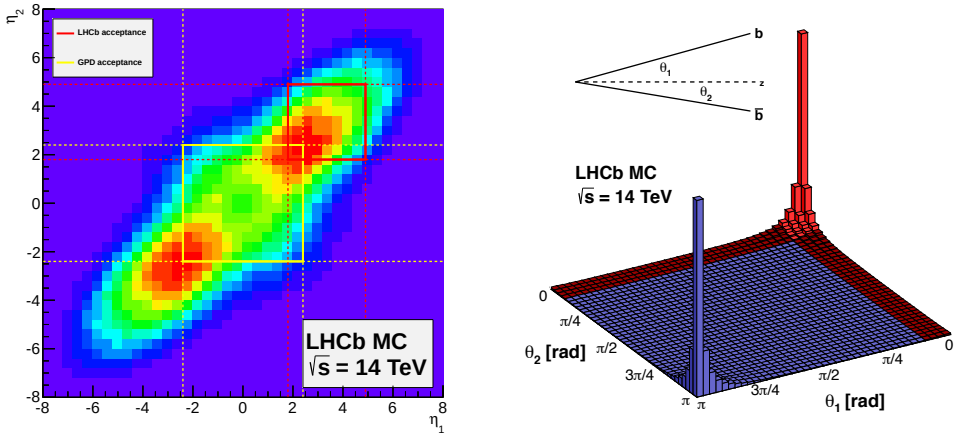

Figure 3.4: Left: The pseudorapidity of the $b\bar{b}$ quark pairs generated in 14 TeV centre of mass energy proton-proton collisions with PYTHIA 8. The red square region shows the acceptance covered by LHCb while the yellow square the equivalent for the general purpose LHC experiments. Right: The polar distribution of $b\bar{b}$ quark pairs with respect to the beamline. The red shaded region shows the area included within the LHCb acceptance. Figures are taken from [139].

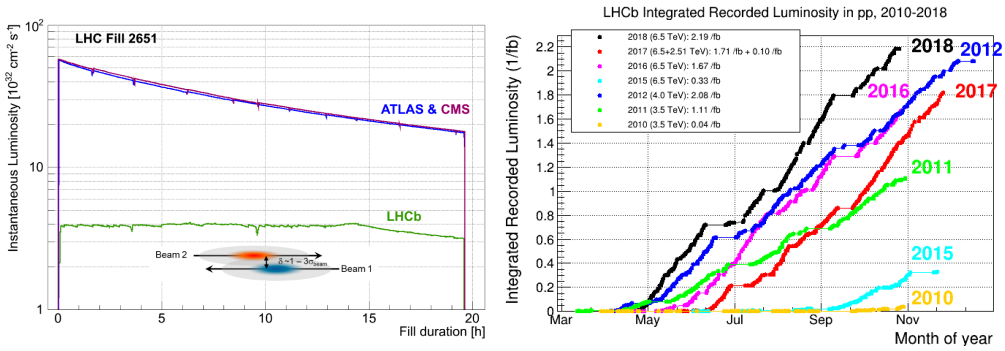

Figure 3.5: (Left) the instantaneous luminosity recorded by LHCb, ATLAS and CMS during a fill [128] and (Right) Integrated luminosity recorded by the LHCb experiment from Run 1 and Run 2 of the LHC [139].

To ensure that the instantaneous luminosity does not become too large, the focus and lateral separation of the beams at the crossing point is carefully adjusted over the course of an LHC fill. As well as preventing the luminosity from becoming too large, this levelling scheme is also used to achieve the opposite effect and compensate for the fact that the longer a fill goes on, the more the instantaneous luminosity reduces due to the loss of protons following beam collisions. The luminosity levelling scheme allows LHCb to maintain this balance between maintaining a high enough luminosity to record sufficient data to perform physics analyses, whilst ensuring

that this luminosity does not become too large as to make vertex reconstruction untenable. A comparison between the instantaneous luminosity profiles of the general purpose detectors and LHCb can be seen in Figure 3.5. It should be noted that to date, during data-taking runs, ATLAS, CMS and LHCb have all operated at luminosities much higher than they were initially designed for. For example, during Run 1, LHCb's operational luminosity was $\mathcal{L} = 4 \times 10^{32}$ cm^{-2}s^{-1} -- twice the initial design value.

The amount of data recorded by LHCb over the course of data taking is quantified in terms of the integrated luminosity, L. As shown in Figure 3.5, LHCb recorded a total integrated luminosity of 3.1 fb^{-1} of data in 2011+2012, 1.9 fb^{-1} of data in 2015+2016 and 3.8 fb^{-1} of data in 2017+2018. It is this full dataset which is used within the analysis presented within this thesis.

3.3 Particle Tracking

The LHCb tracking system is responsible for the reconstruction of the trajectories of charged particles that pass through the detector. The system is comprised of a VErtex LOcator (VELO) that surrounds the interaction region, a warm dipole magnet, the Tracker Turicensis (TT) upstream of the magnet and, finally, a series of three additional tracking stations downstream of the magnet known as T1, T2 and T3. The spatial information which describes where a particle interacts with an element of the tracking system is referred to as a particle *hit* and the LHCb tracking system takes the series of particle hits arising in the different sensors to estimate the particle trajectory. This process is extremely precise within the interaction region to allow for the separation and classification of the PVs and SVs. Additionally, this fine resolution is necessary for the measurement of several different physical observables. For example, B_s^0 mesons oscillate into a $\overline{B}_s^0$ meson and vice versa at a frequency of 17.7ps^{-1}. Under the assumption that the B_s^0 travels at 99% of the speed of light, the tracking system would need to have a spatial resolution of ~ 0.1 mm or better to resolve the two states. The excellent spatial resolution provided by the tracking system and in particular from the VELO sub-detector is complemented with precise measurements of the track momenta from a high bending power dipole magnet with the precision in the momentum measurement resulting in a good resolution in reconstructed invariant masses.

Particle tracks at LHCb are classified depending on which sub-detectors they pass through as is illustrated in the schematic diagram of Figure 3.6. Charged particles which only leave hits within the VELO system are known as VELO tracks and tracks which pass through both the VELO and TT tracker before being swept out of the acceptance region by the magnet are known as Upstream tracks. Tracks which pass through the TT and the T1-3 stations are known as Downstream tracks while those which only pass through the T1-3 stations are known as T tracks. Finally, tracks which pass through all of the elements of the tracking system are known as Long

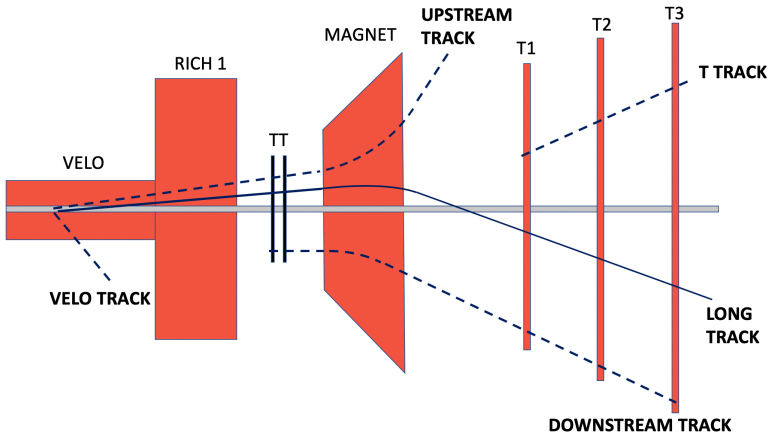

Figure 3.6: The different types of track which may pass through the LHCb tracking system. Tracks which only register hits within the VELO are classified as VELO tracks and tracks which register hits only in the VELO and TT tracker are known as Upstream tracks. Tracks which register hits only in the TT and T1-3 stations are known as Downstream tracks and those only passing through the T1-3 stations are known as T tracks. Tracks which pass through all elements of the tracking system are known as Long tracks.

tracks. The different classes of tracks are often separated from one another as each class can be used to study different phenomena. For example, T tracks can be used to help identify decays which originate from a K^0_S meson which fly a long way through the detector and only tend to decay by the time they have reached the T1-3 stations. However, due to the lack of momentum measurements, T tracks are not useful for most physics analyses. In addition, VELO tracks can be used in the identification of PVs within an event. Long tracks have the most precisely measured momenta as they are reconstructed from hits from both the VELO and all of the tracking stations. As a result they are the most useful for physics analyses and they are the tracks of interest within this thesis.

3.3.1 VELO

The VELO is made up of two sets of 21 semi-circular silicon strip detectors which measure particle hits using an (r,ϕ,z) polar coordinate system. Each VELO module is split into an R-sensor which measures the radial distance of a particle hit from the beam and a ϕ-sensor which measures the hit's azimuthal position. The position of the hit in the z direction is determined from the VELO module which detects the hit signal. The positioning of the VELO modules in the z direction is such that a track which falls into LHCb's acceptance must create a hit in at least four VELO modules. Therefore, the modules which are further downstream are separated

from one another by a larger distance than those closer to the interaction point. The minimum separation between VELO modules is set to be 3.5cm. A schematic of the layout of the VELO can be seen in Figure 3.7 while a diagram of an individual module can be seen in Figure 3.8.

The arrangement of the silicon strips on the sensors are as follows. The R sensor silicon strips are arranged in a concentric fashion with the pitch of the strips increasing linearly from a value of 38μm on the inside edge of the sensor to a value of 102μm on the outside edge. In order to ensure that a single strip does not suffer from an over-capacitance, the sensor is split into four 45° sections with 512 strips present in each section. Without this separation, a large number of hits in a single sensor can result in the individual hits becoming unresolvable from one another which leads to a large drop in strip efficiency. The ϕ strips are arranged radially with the pitch increasing linearly from inner edge to outer edge. In order to make sure that the pitch does not become too large at the outer edge, which would reduce the sensitivity of the module, the sensor is split into two regions. In the inner section, which contains 683 strips, the pitch is 38μm at the inner edge and 78μm at the upper edge. At the beginning of the second section, which contains 1365 strips, the pitch resets to 39μm before increasing to a maximum pitch of 97μm at the outer edge of the sensor. There is a slight skew of 20° in the inner region and 10° in the outer region between the R and ϕ strips meaning that the two sets of strips are not perfectly orthogonal and this slight skew introduces a stereo angle. When consecutive VELO modules are placed next to one another they have an opposite skew and the total polar resolution of the VELO is improved. Finally, in addition to the primary VELO modules, and in order to measure the pile-up at the interaction region, each set of VELO modules includes two additional sensors, known as VETO stations which are upstream of the other modules. These additional sensors are only comprised of R-sensors.

During the injection and ramping stages of an LHC fill each set of detectors is placed, either side of the beam, at a distance of 30mm from the beam to prevent unnecessary radiation damage. Once the LHC beam is stable and is ready for data taking the VELO modules are moved into position with the active region of the detector 8.2mm away from the beam. Having a measurement point close to the beam improves the IP resolution, which results in an improved vertex separation capability. Due to small variations in the position of the LHC beam from one fill to another it is also necessary for the VELO to be re-centred around the beam slightly differently each time it is closed. When closed the VELO modules slightly overlap one another by approximately 1.5mm. This is a design tolerance intended to make sure that there is a full coverage of the possible hit region.

Both the modules themselves and the corresponding readout electronics are placed inside an evacuated vessel. Inside the vessel the VELO modules are surrounded by a radio frequency (RF) foil which is made of 300μm thick aluminium. The RF foil is designed to shield the sensitive

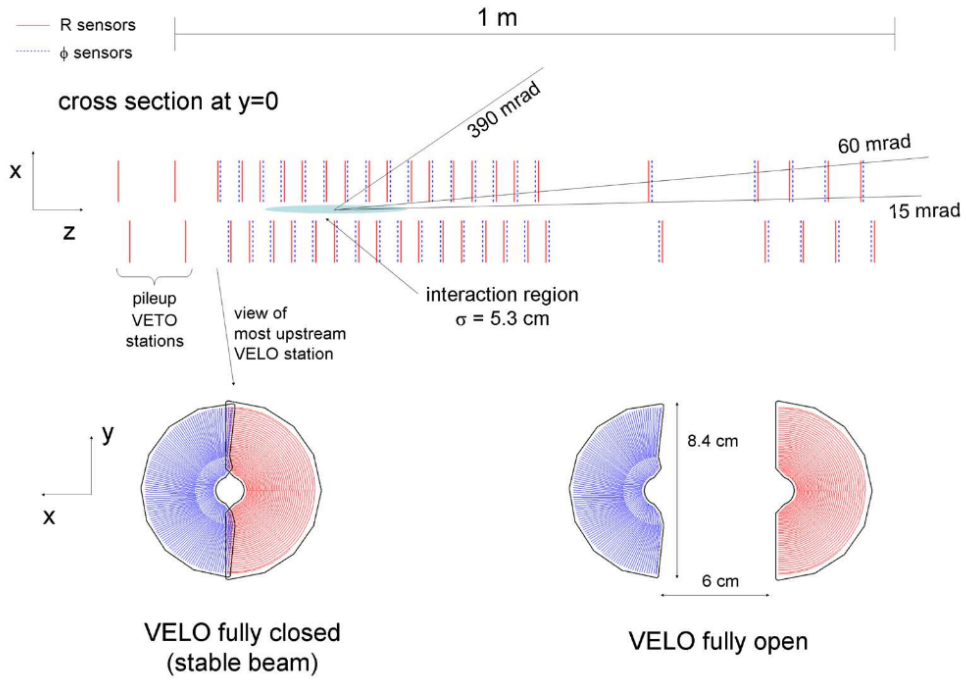

Figure 3.7: Distribution of VELO modules along the LHC beamline including the VETO stations upstream and 42 standard modules downstream [127]. The lower half of the figure shows the configuration of the VELO module in both the closed (left) and retracted (right) state.

VELO modules from any RF signals which originate from the beam and which would interfere with the VELO readout whilst also ensuring that the VELO modules do not contaminate the separate LHC beam pipe vacuum.

3.3.2 Tracker Turicensis

The Tracker Turicensis (TT) is a 150cm$\times$130cm silicon strip detector that covers the entire LHCb acceptance. It is located after RICH 1 and before the LHCb dipole magnet and its function is two-fold. Firstly, the TT provides a measurement point before the magnet which allows for the reconstruction of long lived particles, such as a K^0_S or Λ^0, which do not decay inside the VELO. Secondly, the TT enables the reconstruction of soft, low momentum, particles which are deflected by the magnet such that they never reach the T1-3 Tracking stations further downstream. The TT, which can be seen in Figure 3.9, is made up of four layers which are arranged into two pairs, with each pair separated by approximately 27cm in the z direction. The first and last layers are arranged parallel to one another in a vertical direction. The inner two layers are skewed by $+5^\circ$ and -5° respectively to the vertical giving the TT tracker a so-called (x-u-v-x) structure where x

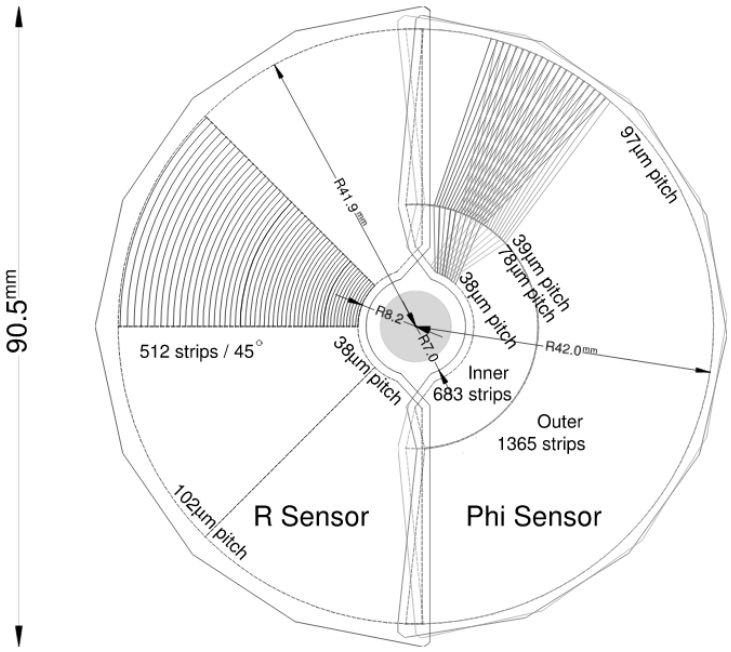

Figure 3.8: Schematic of a single VELO module [127]. The left sensor corresponds to the R sensor with concentric silicon strips while the right the ϕ sensor with radial strips. The ϕ half of the figure is comprised of two superimposed, consecutive VELO modules in order to highlight the skew between strips on adjacent modules.

represents the two end layers and u and v the two inner layers. This skew is included to introduce a stereo angle, similarly as is done for consecutive VELO modules. Each layer of the TT is made from a half-module above the beampipe and another half-module below the beampipe and each half-module comprises of a column of seven silicon strip sensors each with 512 readout strips where the pitch between adjacent strips is 183μm. A single half module is also divided into 3 sectors named *L*, *M* and *K* where the different sectors have separate readout channels and different numbers of sensors. The *L* sector covers the area of the module furthest from the beam and comprises of 4 sensors. The *M* sector is then placed closer to the beam and comprises of 2 sensors whilst the *K* sector is placed closest to the beam and consists of a single sensor. By providing dedicated readout channels for the sectors the possibility of electrical over-capacitance is guarded against. The area closest to the beam, given that the area is expected to have the most number of tracks passing through it, requires it's own channel. The outer sectors receive fewer tracks and so sensors can be grouped together safely without risking signal overloading.

At the end of the modules, outside of the detector acceptance, lie the readout electronics, structural supports and a cooling system. The cooling system is designed to keep the TT at

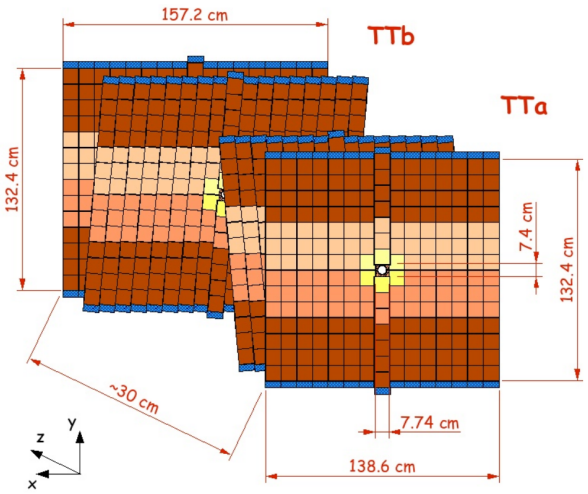

Figure 3.9: The Tracker Turicensis (TT) with the four layers visible in the (x-u-v-x) orientation [140]. The blue region houses the hybrid readout electronics, the dark orange regions the *L* sectors, the lighter orange sectors the *M* sectors and the yellow regions the *K* sectors.

temperatures less than 5° C. The four layers which make up the TT are housed inside a light-tight, thermally and electrically insulated volume. The volume is continuously flushed with nitrogen gas to ensure that no condensation forms on any components of the detector.

3.3.3 Inner Tracker

The three tracking stations T1, T2 and T3 are split into two different regions known as the Inner Tracker (IT) and the Outer Tracker (OT) with these two regions highlighted in Figure 3.10. The IT covers the acceptance region where the track multiplicity is largest and it's function is to help reconstruct those tracks which pass through the magnet very close to the beamline. Approximately 20% of charged particle tracks which pass through the tracking stations do so via the IT. The structure of the IT is very similar to that of the TT; the IT is comprised of four layers of silicon strip detectors arranged in the (x-u-v-x) configuration with the inner layers having a $+5^{\circ}$ and -5° skew. The IT section of each of the three tracking stations is comprised of 4 light-tight boxes which house the silicon sensors and which surround the beam pipe. The sensors above and below the beam pipe comprise of a single layer of 7 silicon sensors, while the sensors either side of the beam pipe comprise of two layers of 7 sensors. Each sensor comprises of 384 silicon strips with a strip pitch of 196μm. Like the TT tracker, the IT is cooled to temperatures below 5° C and constantly flushed with nitrogen to prevent condensation. However, unlike the TT, the readout electronics and the cooling system are included within the detector acceptance. A schematic of an IT module can be seen in Figure 3.11.

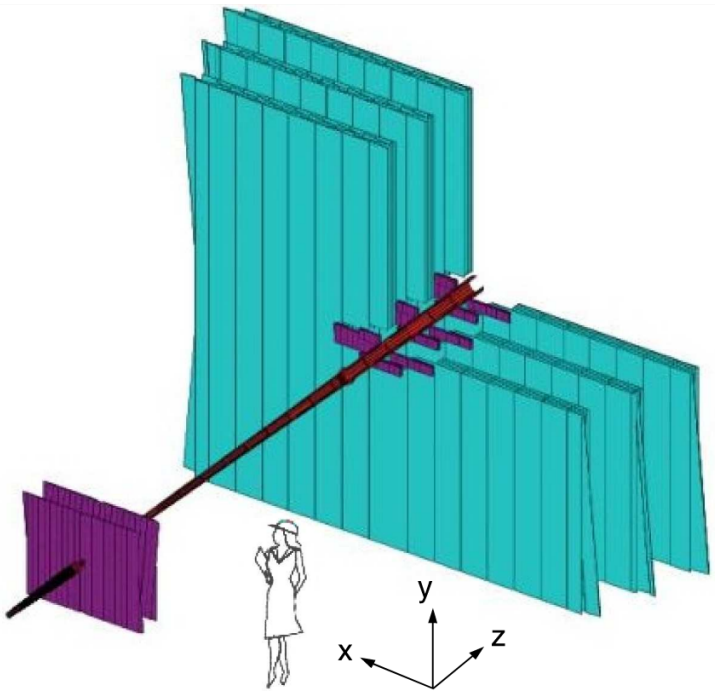

Figure 3.10: An angular view of the TT and the T1-T3 [127]. The purple regions of the downstream tracking stations show the IT and the turquoise sections the OT.

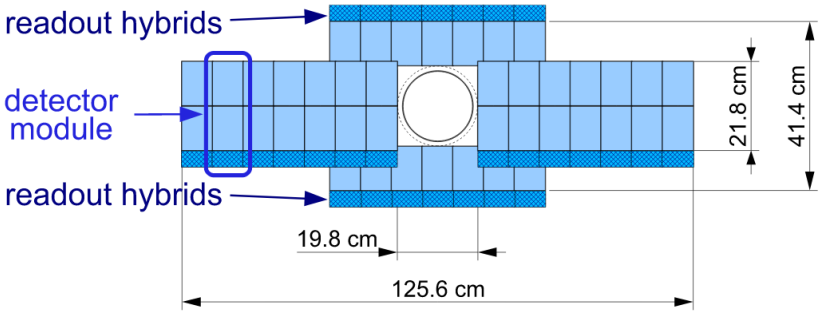

Figure 3.11: Layout of an IT module from the second IT station [127].

3.3.4 Outer Tracker

The OT, which accounts for 98.7% of the area of the tracking stations and which is shown in Figure 3.12, reconstructs the tracks traversing the acceptance region not covered by the IT. Unlike the TT and IT and due to it's large, 30m^2 area, the OT does not provide tracking capability via silicon strip technology. Instead, each OT station comprises of four modules of straw drift tubes, with the modules arranged in the same (x-u-v-x) configuration as used by both the TT and IT.

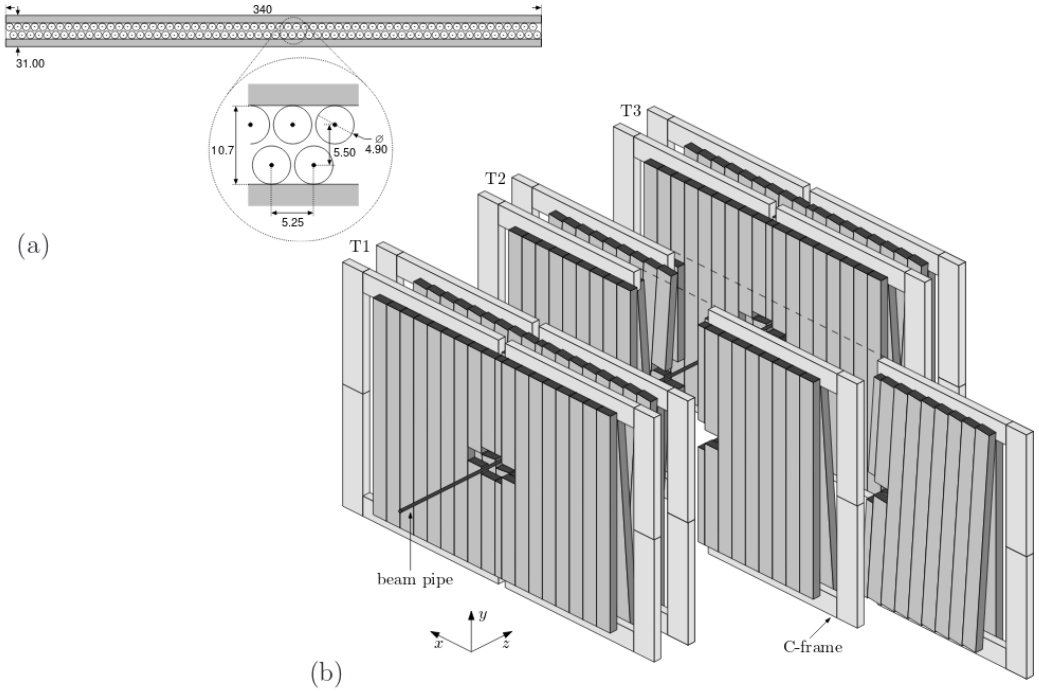

Figure 3.12: The OT detector [141], showing a) a cross section of a single OT layer with the staggered rows of drift tubes and their dimensions visible and b) the overall layout of the OT with each tracking station being comprised of four modules which each have two OT layers. The figure shows the re-tractability of the OT modules.

Each module contains two staggered layers of 64 drift tubes, with each tube having an inner diameter of 4.9mm and a length of 2.4m. The tubes themselves are made of two layers of foil with the inner, carbon-doped polyamide layer acting as a cathode and a gold-plated tungsten anode running down the tube centre. The drift tubes are filled with a gas containing Ar, CO_2 and O_2 in the ratio 70%:28.5%:2.5% and when a charged particle track passes through the tube the gas is ionised. The ionised charge then drifts to the anode in the centre of the drift tube via the application of an electrical current. The measured drift time then provides information on the location of the ionisation. The gas composition was chosen to ensure that drift times would be kept below 50ns and that the drift resolution would be kept below 200μm.

3.3.5 Dipole Magnet

In order to measure the momenta of different particle tracks passing through the detector with sufficient precision, LHCb makes use of a warm dipole magnet with a bending power of 4Tm, which corresponds to the integrated magnetic field experienced by a track 10m in length. The

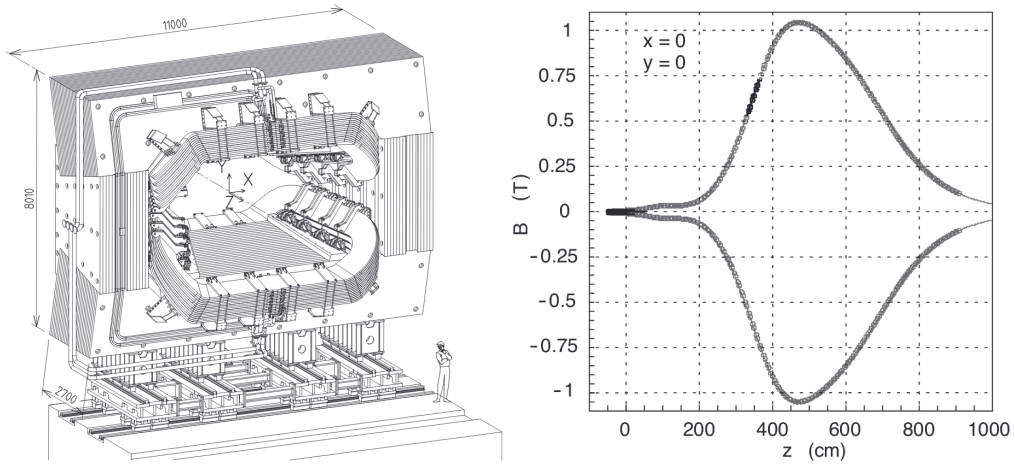

Figure 3.13: A schematic of the LHCb warm dipole magnet (left) [127]. The coordinate system is defined by a set of axes in the middle of the schematic. The magnetic field strength of the LHCb dipole magnet as a function of z, the beam axis, as measured by Hall probes (right) [142].

magnet produces a magnetic field parallel to the y direction and, in conjunction with the tracking detectors, is able to achieve a momentum precision of $\delta p/p < 5 \times 10^{-3}$ for particles with a momentum of less than $100\,\mathrm{GeV}/c$. The magnetic field covers the full LHCb acceptance and therefore can provide a momentum measurement for most particle tracks. However, the magnet cannot provide good momentum measurements for the Upstream tracks which typically have momenta $< 1\,\mathrm{GeV}/c$ and which are swept out of the detector acceptance. Additionally, the magnet cannot provide a momentum measurement for T tracks.

The magnet is comprised of two identical coils placed symmetrically about the beam axis with each coil made of fifteen aluminium layers with the layers enclosed inside an iron yoke. In order to determine the momentum of a particle track the magnetic field strength must be known precisely and the profile of the magnetic field must be well understood. Therefore, the magnetic field profile as a function of z is measured with Hall probes with the profile seen in Figure 3.13. It can be seen that the strength of the magnetic field drops off quickly as you move away from the magnet itself. The strength of the magnetic field present in the VELO region is very small meaning that the tracks inside the VELO are straight and that the VELO resolution is not compromised. The magnetic field in the regions containing the RICH detectors (discussed further in Section 5.5) is larger which could provide a problem for the RICH system's sensitive photon detectors. As a result, the RICH detectors are enclosed with iron shielding to prevent any interference from the magnetic field.

Over the course of a data taking period the polarity of the magnet is flipped at regular intervals such that the datasets recorded by LHCb can be approximately divided equally into

a “Magnet Up” and “Magnet Down” state. This is done to ensure that there are no systematic uncertainties which enter LHCb analyses due to some asymmetry in the detector which is especially important for analyses which study CP asymmetries.

3.3.6 Tracking and Vertexing Performance

The process of particle track reconstruction starts from the identification of particle hits in the various tracking sub-detectors. Straight hit trajectories in both the VELO and the TT are identified as these are regions where the magnetic field strength is relatively small. These trajectories are then matched to hits in the tracking stations and any duplicate tracks are removed. Then, using a Kalman filter, the tracks are fit to improve the estimated track parameters. This process accounts for multiple scatterings and loss of track energy in the detector. It is possible that tracks are reconstructed which do not correspond to any trajectory in the detector. These are referred to as “Ghost” tracks and these tend to originate from an incorrect matching of hits in the different tracking detectors. Ghost tracks can be removed using multivariate analysis techniques which take track kinematics, trajectories and particle hits as inputs.

The vertexing resolution is measured from data by taking vertices which have a large number of tracks. Then, the tracks are randomly split into two groups with the fit repeated for the two groups separately to obtain two separate measurements of the vertex position. The resulting difference in vertex position between the two groups is then measured [128]. The variation in the reconstructed vertex position can then be fitted with a distribution whose width can be used to infer the PV resolution. It can be seen in Figure 3.14 that for a typical event with 25 reconstructed particle tracks the resolution of the primary vertex is $13 \times 13 \times 71 \mu m$ in $x - y - z$. Additionally, the vertexing resolution is seen to improve in the z co-ordinate in Run 2. Within Run 2 only VELO tracks were used to reconstruct the PVs whereas in Run 1 Long tracks were also used. In Run 2 this led to a more consistent set of PV positions throughout the analysis chain and smaller systematic effects.

Another common metric to describe the performance of the VELO is the measurement of the Impact Parameter (IP) resolution for a given track. The IP of a track is defined as the distance of closest approach between the track and the PV. It is a useful parameter for heavy flavour decays because the displaced SV will result in a larger value of the IP compared to background events. The IP resolution is measured in a similar way to the vertexing resolution [128]. Firstly PVs are selected and the IP of a random track is measured. Then, the random track is removed from the vertex fit and the fit is repeated and the IP recalculated. After repeating the exercise, the distribution of differences in IP values is taken as the IP resolution. A measure of the IP resolution as a function of p_{T} can be seen in Figure 3.15 where the plot shows that superior resolutions are

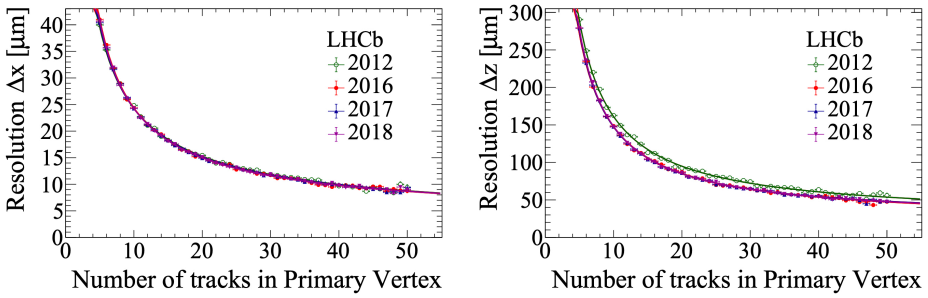

Figure 3.14: The resolution of the PV as a function of number of tracks in the event for both Run 1 and Run 2 in both the x direction (left) and the z direction (right) [143]. The resolution in the x direction is the same as the resolution in the y direction due to the VELO geometry. The improvement seen in the z resolution in Run 2 is due to the PV algorithm only using VELO tracks leading to a reduction in systematic effects.

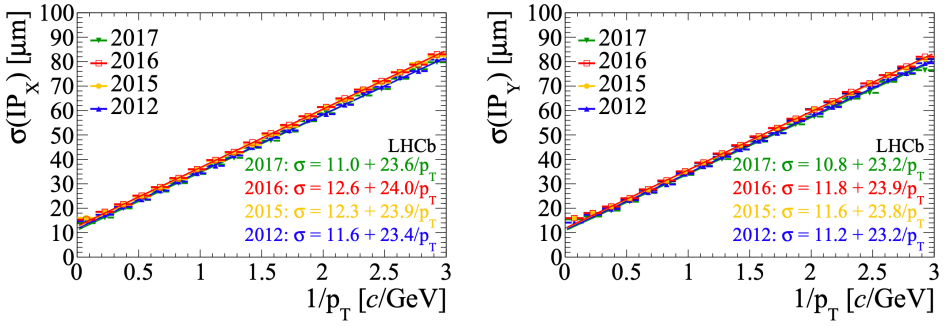

Figure 3.15: The IP resolution as a function of p_{T} for Run 1 and Run 2 for both the x direction (left) and the y direction (right) [143].

achieved at higher transverse momenta.

The performance of the TT and IT trackers is studied by measuring the tracking resolution using an experimentally clean sample of $J/\psi \to \mu^+\mu^-$ decays [128]. The resolution can be estimated by comparing the position of the hit in the detector to the position estimated based on the other hits on the track. The distribution of the difference between these two positions provides the resolution. For both the TT and IT resolutions of $\sim 50\mu m$ have been achieved. This resolution varies slightly across the TT tracker and is seen to improve towards the edge of the detector due to a larger track angle resulting in more charge sharing between strips. This effect can be seen in Figure 3.16.

The performance of the OT is primarily dependent on the drift time resolution. In Run 1 drift time resolutions of less than 3ns were achieved alongside a single-hit efficiency per module of over 99%. In Run 2 this resolution was improved by approximately 20% due to improvements

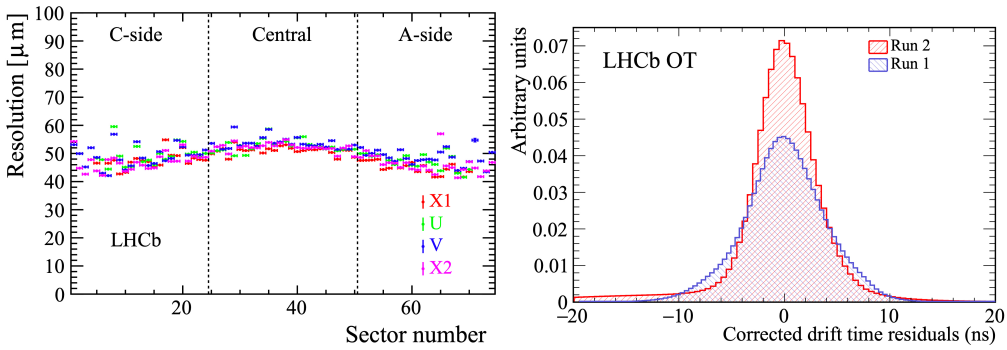

Figure 3.16: The left plot shows the resolution of particle hits across the TT [128]. The different modules in the (x-u-v-x) configuration are labelled and a slight improvement in the resolution can be seen at the edges of the tracker. The right figure shows the distribution of the drift time residuals for both Run 1 and Run 2 in the OT [144]. The approximate 20% improvement in resolution in Run 2 can be seen.

in the calibration procedure, as can be seen in Figure 3.16.

The total tracking efficiency for reconstructed tracks is evaluated using a tag and probe method with a sample of $J/\psi \to \mu^+\mu^-$ decays. The tag and probe method involves taking the first muon track, the tag, and fully reconstructing it to select events. The second muon track, the probe, is then only partially reconstructed. The fraction of probe muons which are successfully matched to a long track in the event then defines the tracking efficiency. As can be seen from Figure 3.17 the tracking efficiency is normally very high and is in excess of 96% for particle tracks with momenta between $5 < p < 200\,\text{GeV}/c$. When trying to replicate the tracking efficiency with LHCb Monte Carlo (MC) simulation, generally, the simulation describes the tracking efficiency well. However, small discrepancies are corrected for by weighting the MC in bins of track momenta and pseudorapidity which the tracking efficiency is sensitive to.

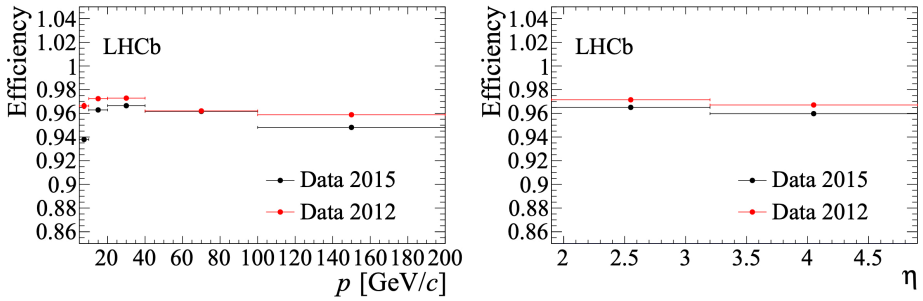

Figure 3.17: The efficiency in tracking reconstruction as a function of both the track momentum (left) and the pseudorapidity (right). The efficiencies are shown for 2012 and 2015 with similar behaviour seen for the rest of both Run 1 and Run 2 [143].

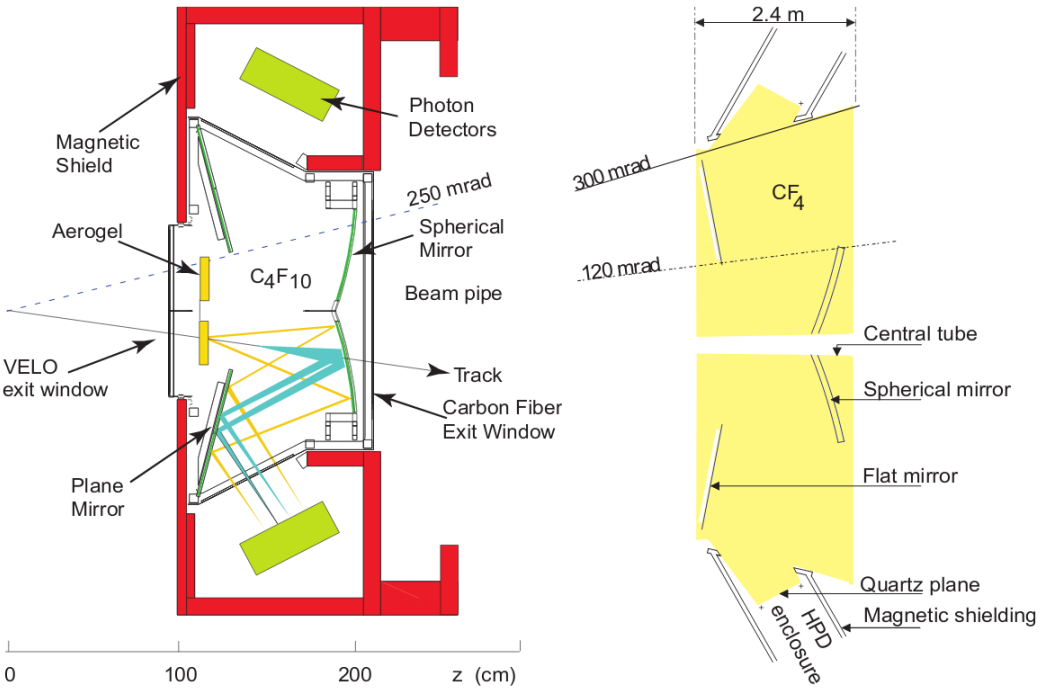

Figure 3.18: A cross-section of RICH 1 (left) in the side view, y-z plane. Also a top-down, x-z plane, view of RICH 2 (right). The aerogel plates were removed after Run 1 due to poor performance [128].

3.4 Particle Identification

The ability to identify different species of particles is crucial for most heavy flavour physics analyses. In particular, the ability to distinguish between pions, kaons and protons is especially important for LHCb as these are the hadronic final state particles which exist in the finite length covered by the detector.

The LHCb detector has dedicated sub-detectors which are designed for the specific task of performing Particle IDentification (PID). These include the Ring Imaging Cherenkov Detectors (RICH) RICH 1 and RICH 2 as well as a muon detector system. In addition, information from the calorimeter system (see Section 3.4.2) is combined with information from these sub-systems so that the LHCb detector can also separate out electrons, photons and neutral pions.

3.4.1 RICH Detectors

Hadronic PID information is provided by two RICH detectors, RICH 1 and RICH 2, seen in Figure 3.18. These exploit the Cherenkov effect, whereby a charged particle travelling through a material at a speed quicker than light travels through the medium results in the emission of photons in a cone. The angle of the cone produced depends both on the refractive index of the material and the velocity of the particle in question and can be expressed as

$$\cos(\theta_c) = \frac{1}{\beta n} \tag{3.4}$$

where θ_c is the Cherenkov angle, β is the velocity as a fraction of the speed of light and n is the refractive index. The idea behind the RICH system is that it is possible to exploit this by measuring the angle and then combining the angle with the momentum measurements from the tracking system to infer the mass and hence the species of the particle. The different hypotheses as a function of Cherenkov angle can be seen in Figure 3.19. In practice, hits in the RICH detectors produce hit patterns which are compared to images of hits from different mass hypotheses in a likelihood calculation.

The first RICH detector, RICH 1, is located downstream of the VELO and upstream of the TT. RICH 1 covers the full angular acceptance of LHCb and is designed to provide PID for low momentum particles; defined as those particles with momenta in the range 2-60 GeV/c. In Run 1 of data taking the Cherenkov medium used for RICH 1 included aerogel tiles which were designed to cover the momentum range 2-10 GeV/c and perfluorobutane (C_4F_{10}) designed to cover the momentum range 10-60 GeV/c. It was observed in Run 1 that the ability of the aerogel to improve the PID performance was poorer than predicted. As a result, in Run 2, the aerogel was removed from RICH 1 to reduce the amount of material in the detector. This did not result in a noticeable drop in PID performance although did speed up the RICH reconstruction process [145]. RICH 2 is located further downstream; between the T1-3 tracking stations and the calorimeters. This location is chosen to reduce the amount of material which tracks have to travel through before arriving at the tracking stations. The Cherenkov material of RICH 2 is tetrafluoromethane (CF_4) and is chosen to provide PID for higher momentum particles with momenta between 15-100 GeV/c. The angular acceptance of RICH 2 is 15-120mrad, instead of the full LHCb angular acceptance. This smaller acceptance covers the range in which high-momentum particles are produced and is chosen because lower momentum particles will not reach RICH 2 as they are swept out by the magnet.

The Cherenkov photons emitted by charged particles, in both RICH 1 and RICH 2, are directed and focused, using both flat and spherical mirrors, onto a plane of hybrid photodetectors

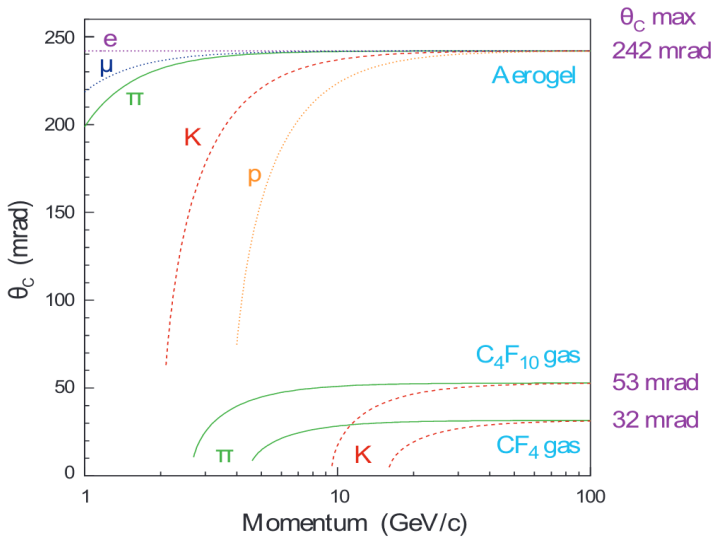

Figure 3.19: The dependence of the Cherenkov angle on particle momentum for different particle species. The radiative mediums used by both RICH 1 and RICH 2 and labelled [127].

(HPDs). The HPDs, of which there are 196 in RICH 1 and 288 in RICH 2, are vacuum photon detectors. When incident photons hit the HPD cathode they release photoelectrons which are then accelerated by an applied voltage onto a pixelated silicon detector. The HPDs are enclosed inside an external iron shield, outside of the detector acceptance, in order to shield them from the dipole magnet's strong field. This is necessary as in the presence of a magnetic field, photoelectrons can be swept out and lost from the active area of the HPD which dramatically lowers its efficiency.

3.4.2 Calorimetry

The calorimeter system comprises of four sub detectors. These are the Scintillator Pad Detector (SPD), the Pre-shower Detector (PS), the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). The calorimetry system both measures the energy of particles and is crucial in the identification of photons, electrons and neutral pions. Information from the calorimeters is also used within the LHCb hardware trigger to select physics events of interest before saving them to disk (discussed more in Section 3.5.) Specifically, the calorimeters are tasked with identifying those events which have a high transverse energy, E_T, which are likely to stem from the decay of a beauty hadron.

All four sub detectors of the calorimetry system work on the same principle. As particles fly through the calorimeter they interact with the material of the calorimeter and produce cascading particle showers. In the case of an electromagnetic shower this process takes place via pair

production, Compton scattering and bremsstrahlung radiation. Alternatively, within a hadronic shower, the cascade is created via inelastic processes which produce copious amounts of pions and further electromagnetic showers. The photons from the particle shower are transported, via wavelength-shifting fibres, to arrays of photomultiplier tubes where the intensity of light provides a proxy for the energy of the incoming particle. The material used in the calorimetry system is a doped polystyrene and layers of this material are separated from one another by a layer of dense absorber material where the showers originate from.

Another common feature of the calorimetry sub-detectors is that they are all separated into different sections to process different regions of the LHCb acceptance where there is a different occupancy. The PS, SPD and ECAL have three sections which include, as can be seen in Figure 3.20, an inner section (where the occupancy is expected to be highest), a middle and finally, an outer section. The HCAL has only two regions; an inner and an outer region.

The PS and SPD detectors are immediately upstream of the ECAL and downstream of the first muon detector and are located before the location of any particle showers. These are single layered scintillator pad detectors which are separated from one another by 15mm of lead. The SPD's function is to determine whether or not the incident particle is electrically charged as charged particles are more likely to deposit energy in the SPD layer whereas neutral particles are more likely to proceed and deposit their energy in the ECAL. The PS detector then complements this by distinguishing electrons from hadrons as electrons are more likely to begin showering in the separating lead absorber. Therefore, when combining the SPD and PS output with information from the ECAL a distinction can be made between photons and electrons, despite the fact that these different particle species all produce similar electron-photon showers in the ECAL itself.

The ECAL, which has a design energy resolution of $\sigma_E/E = 10\%/\sqrt{E[GeV]}$, comprises of alternating layers of 2mm thick lead and 4mm thick scintillating tiles with the layers arranged transverse to the beam. In total, there are 66 layers of lead and scintillating tiles. The

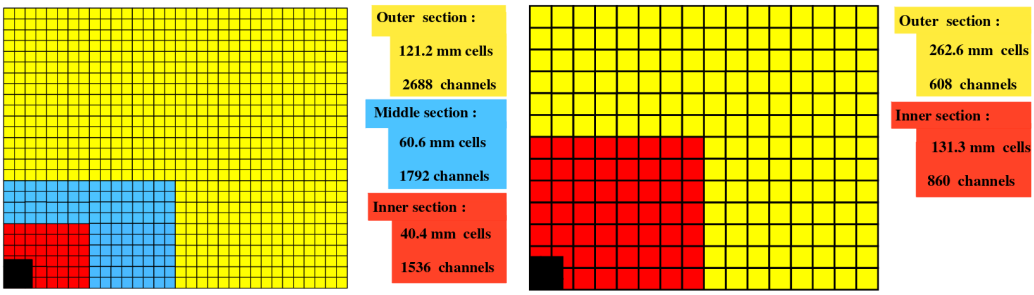

Figure 3.20: The different scintillating regions of the SPD, PS and ECAL (left) and the HCAL (right) with the position of the beam pipe in black [128].

reason for choosing this thickness and number of layers is that, together, this corresponds to 25 radiation lengths for photons and electrons which means that the showers should be completely contained within the ECAL.

Finally, the HCAL is made of 216 alternating layers of 3, 7mm thick, iron plates and 3, 3mm thick, scintillator tiles arranged parallel to the beam axis. Due to space constraints in the cavern housing LHCb, the depth of the HCAL is limited to 1.6m which corresponds to 5.6 nuclear interaction lengths. As a result, high energy showers are not entirely contained by the HCAL but this does not impact the HCAL's ability to identify high E_T particles which is it's primary function. Another difference between the ECAL and the HCAL is that the HCAL readout cells are larger than those in the ECAL due to the different structure of hadronic showers compared to electromagnetic showers. The different cell structure and finite length of the calorimeter results in an energy resolution of $\sigma_E/E = 69\%/\sqrt{E[GeV]}$.

3.4.3 Muon Detection

The muon system enables the detection of muons which result from many of the rare, FCNC decays of interest, as was mentioned in the Chapter 2. In addition, and similarly to the calorimeter system, p_{T} information from the muon detectors is included in the trigger system. In total, there are five separate muon stations, referred to as M1-M5, and these are shown in Figure 3.21. Due to the ability of muons to pass through most materials with minimal interaction, the muon stations are positioned at the most downstream extent of the LHCb detector and muons are the only charged particle which reaches them. The exception to this is M1 which is placed just upstream of the SPD and PS detectors in order to provide an improved p_{T} measurement for use in the trigger, before the muon passes through the dense material of the calorimetry system.

The muon stations are made from four concentric regions known as R1-R4 with the division made in order to ensure that the occupancy of each section is approximately the same. Apart from section R1 of station M1 (which is made of triple-Gas Electron Multipliers), the muon stations are made from staggered layers of Multi-Wire Proportional Counters (MWPCs). The MWPCs are comprised of a pair of cathode plates separated by 5mm which are filled with a gas mixture of Ar, CO_2 and CF_4 in the ratio 40%:55%:5% respectively with an anode wide running down the middle. The multipliers used in M1 have a higher radiation tolerance than the MWPCs which is appropriate given their position further upstream in the detector. The principle of operation for both the MWPCs is the same as the straw drift tubes in the OT and a time resolution of approximately $3\,\mathrm{ns}$ is achieved in the MWPCs.

The muon stations further downstream, M2-M5 are each separated with iron shielding plates which are 80cm thick. These plates serve to reduce any background from non-muon

particles which have managed to penetrate far downstream. In addition, M1-M3 have a larger granularity than M4 and M5. Stations M1-M3 are used to establish the direction of muon tracks whilst providing a 20% momentum resolution in the bending plane. Meanwhile, M4 and M5, being at the most downstream point of the LHCb detector, are used to identify the most penetrating particles. As a result, the spatial resolution of M4 and M5 is reduced compared to the earlier stations. In order to pass through all of the muon stations a muon is required to have a momentum of at least 6 GeV/c.

3.4.4 PID performance

Ultimately, the PID hypothesis for a particular particle track depends on the information from the entire PID system, the RICH detectors, the calorimeters and the muon detectors. The output from these different subsystems can be combined to form a likelihood that the particle is a specific species of particle that can then be compared to the likelihood that the particle is a pion. An alternative technique, which is now more commonly used, uses a multivariate analysis which takes the likelihoods as an input along with other variables which describe the whole event, such

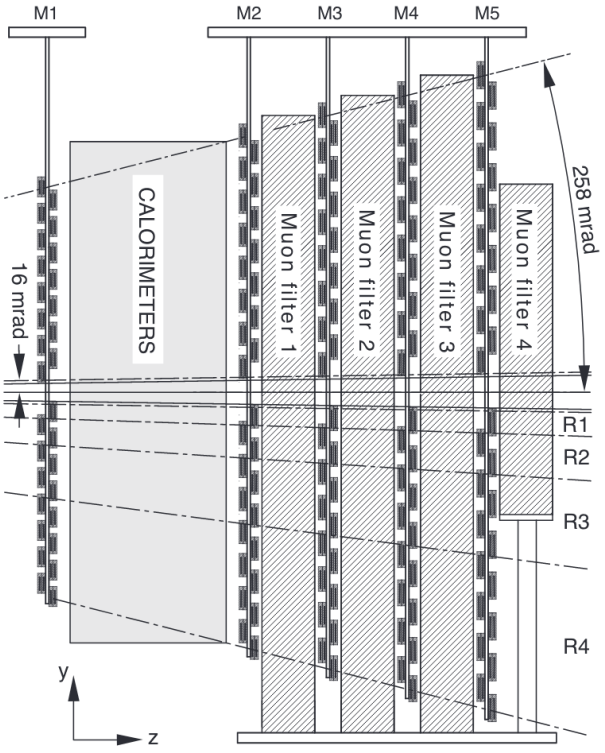

Figure 3.21: A side view of the muon detectors M1-M5 [146].

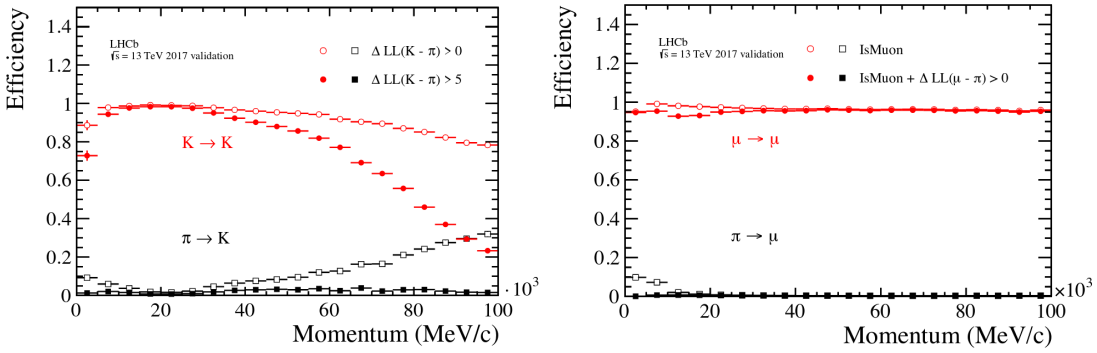

Figure 3.22: Different PID efficiencies for different particle hypotheses as a function of momentum, taken from 2017 Run 2 data. The left plot shows the efficiency of correctly identifying a kaon and mis-identifying a pion as a kaon. The right plot shows the efficiency of correctly identifying a muon as a muon and mis-identifying a pion as a muon [139].

as the event multiplicity. The multivariate analysis then returns the probability that the particle track corresponds to a particular hypothesis and, unlike the bare likelihood values, accounts for correlations between the different input variables.

The PID performance is determined using a data-driven method, from sets of calibration samples. The pion-kaon PID efficiencies are determined from a calibration sample of $D^{*+} \rightarrow (D^0 \rightarrow K^-\pi^+)\pi^+$ decays from which the true particle hypothesis can be obtained by tagging on the pion in the final state which does not originate from the decay of the D^0 and then using the charge of this pion to infer which of the two daughters from the charm meson is a kaon and which a pion. Samples of protons are obtained from $\Lambda \rightarrow p\pi^-$ and $\Lambda_c^+ \rightarrow pK^-\pi^+$ decays. Finally, muon PID efficiencies are obtained from calibration samples of $J/\psi \rightarrow \mu^+\mu^-$ decays, such as $B^+ \rightarrow J/\psi K^+$ whereby one of the final state muons is selected using PID variables and one is used to evaluate the PID efficiency.

The pion-kaon and pion-muon PID performance plots for Run 2 can be seen in Figure 3.22. The PID efficiencies are slightly improved in Run 2 compared to Run 1. This is because the hyperparameters of the multivariate analyses are fine tuned in Run 2 while also being trained on larger data sets, resulting in a better overall separation between PID hypotheses. From Figure 3.22 it can be seen that the PID performance from the RICH system decreases at both high and low momenta. The drop in efficiency at low momenta is attributable to the momentum threshold on the production of Cherenkov radiation while at high momenta is due to saturation of the Cherenkov angle.

3.5 Trigger system

The proton-proton bunch crossing frequency at the LHC is 40MHz. For LHCb, the interaction rate is effectively reduced from this number to around 30MHz due to the presence of abort gaps in the beam and the inelasticity of the proton-proton collisions. Nonetheless, reading out data even at this reduced rate would be impractical, not least because most of this data would not be of interest given the experiment's heavy flavour physics focus. Therefore, LHCb utilises a two-stage trigger system made of a hardware and software stage to select data of interest while reducing the data readout frequency to a manageable readout rate of 1MHz which can then be saved to disk at a rate of 2-5kHz in Run 1 and 12kHz in Run 2 [147, 148].

The hardware based trigger, referred to as the Level Zero (L0) trigger, is responsible for reducing the readout from 30MHz to 1MHz. The L0 trigger selects events with a large E_T deposited in the calorimeter system or a large muon p_{T} detected in the muon system. The information from the tracking and RICH detectors is read out after the L0 trigger stage is passed.

The calorimeter based trigger selects high E_T events by summing together cells in either the ECAL or the HCAL whose E_T passes a specific threshold which changes between data taking years and between different particle species. At L0, there is a condition placed on the number of hits in the SPD detector. This condition exists to make sure that events which would take too long to process at the HLT level are removed. The L0 triggers on muons by selecting the two muons passing through the muon system with the highest p_{T}. The trigger compares either these singular values or the product of the transverse momenta to threshold values, with the trigger then firing if the threshold is reached. The p_{T} threshold differs between data taking periods.

All events which pass L0 are then passed to the Event Filter Farm (EFF). The EFF comprises of several thousand computer nodes and it is this computing farm which then runs the algorithms for the software trigger, known as the High Level Trigger (HLT), which has two stages, HLT1 and HLT2. Before the application of the HLT triggers the EFF builds up the event by combining information from the different LHCb subsystems. This accumulated event information is then used to make the final HLT trigger decision.

HLT1 reduces the data rate to 50-111 kHz using a subset of the data where it performs a partial event reconstruction. This involves the reconstruction of PVs within the VELO sub-detector and the extrapolation of VELO tracks upstream to the TT and tracking stations. Once a vertex has been formed it is possible to calculate the IP, the χ^2 value from the vertex fit and the p_{T} of all of the tracks within the event. If the vertices have a small χ^2, the SVs are displaced from the PVs and the tracks have high p_{T} values then the event is of interest and they are then allowed to pass HLT1.

The second level, HLT2, performs a full event reconstruction on the subset of events

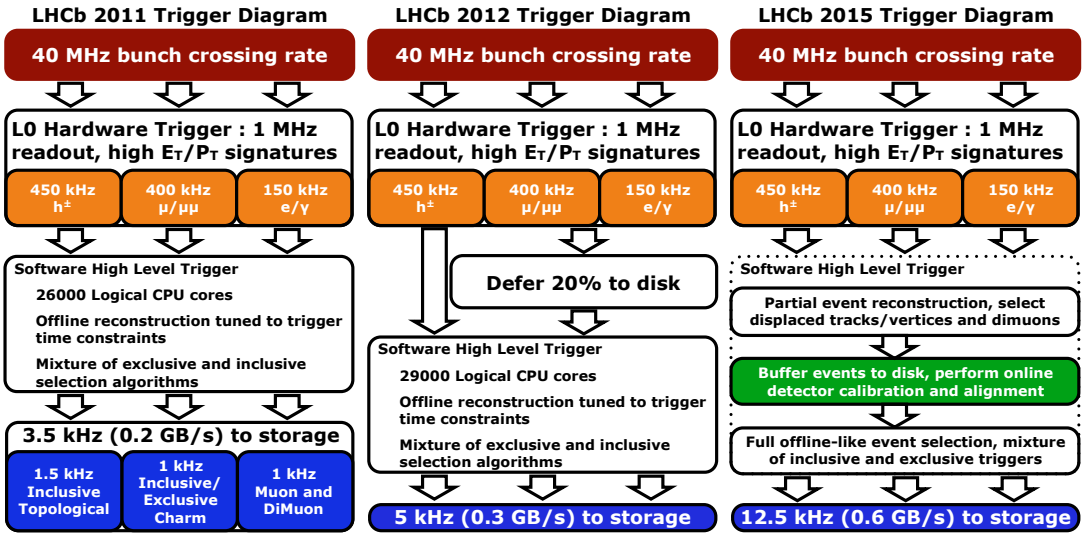

Figure 3.23: The 2011 (left), 2012 (middle) and Run 2 (right) LHCb trigger schemes [149].

which pass HLT1. As HLT2 is exposed to a far lower data input rate, it is able to make trigger decisions using the full event information while making more specific requirements to help search for specific decays which are relevant for physics analyses. HLT2 is associated to two types of triggers known respectively as exclusive trigger lines which are designed to trigger on very specific combinations of final state particles and inclusive trigger lines which trigger on decays which have more general topological features. The topological lines can use multivariate algorithms to help in the trigger decision process and these algorithms take in the kinematics of any tracks associated to the candidate as inputs [150, 151]. Of particular interest in this thesis are the topological trigger lines. These search for candidates with two, three or four charged tracks within the final state that form a displaced vertex that is consistent with originating from a b-hadron decay.

The behaviour and operation of the HLT has been altered over the course of LHC Run 1 and Run 2. At the beginning of Run 1, in 2011, the EFF was not being used when the LHC had finished a fill. In order to avoid wasting computing resources, after 2011, $\sim$20% of events that passed the L0 trigger were sent to local disks with the HLT selection then run on these events between fills. This process resulted in a 25% increase in the HLT processing rate. Furthermore, in Run 2, all events which passed HLT 1 were sent to a buffer. The HLT decision could then be deferred for as long as there was room in the buffer. This meant that the software trigger could be less strict and that a more careful, precise selection could be done at HLT level. Furthermore, due to the presence of the buffer, the detector could be calibrated in real-time during Run 2 and these calibrations could then be made available in the Run 2 trigger.

CHAPTER 4

Analysis Strategy

> *An experiment is a question which science poses to Nature and a measurement is the recording of Nature's answer*
>
> Max Planck,

This chapter provides an introduction to the branching fraction measurement of the rare decays $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$. Specifically, this chapter outlines the measurement strategy used and highlights some of the key aspects and challenges of the analysis.

4.1 Branching Fraction Measurement Strategy

As discussed in Chapter 2, the decays $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ are believed to mainly proceed via the $\rho^0(770)$ and $f_0(980)$ resonances which subsequently decay to $\pi^+\pi^-$. These decays, for which some of the leading order Feynman diagrams can be seen in Figure 4.1, are examples of rare $b \to d$ and $b \to s$ processes. Due to their rare nature, extracting an accurate measurement of a branching fraction for these modes is not trivial. The signal candidates within the dataset are outnumbered by many orders of magnitude by the various sources of backgrounds whose presence within the dataset makes a branching fraction measurement initially impossible. Therefore, before any measurement is made, a series of selection requirements are applied. These requirements involve the removal of candidates from the dataset which fail to pass specific kinematic, topological or particle identification conditions that preferentially select out signal-like candidates. The analysis presented within this thesis uses a single selection for both decay modes. This is due

to the two rare modes sharing a common set of final state particles and subsequently, a similar response to the selection procedure. As was mentioned earlier, a previous analysis from LHCb has already confirmed the existence of the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ mode while falling short of a 5 standard deviation confirmation of the B^0 mode. As a result, the selection presented within this thesis is optimised to make a statistically significant observation of the B^0 mode. The entire selection process is discussed, in detail, in Chapter 5.

After the application of the selection, the branching fractions of the rare decays are determined for the two modes by performing a maximum likelihood fit to the $\pi^+\pi^-\mu^+\mu^-$ invariant mass distribution of the surviving candidates. The fit is made simultaneously to the three datasets used within the analysis which correspond to the 2011+2012, the 2015+2016 and the 2017+2018 datasets. The branching fraction values are free parameters and are allowed to vary during the fitting procedure meaning they can be directly extracted from the fits. The branching fraction parameter values are directly proportional to the signal yields of the two rare modes. Therefore, the underlying functionality of the fits is to determine these yields while the parameters, which relate the yield and branching fraction, are evaluated prior to the mass fit or are taken from previous literature. As only a single fit is performed to provide a measurement for both rare modes it is necessary for the process to be able to separate the B^0 and B_s^0 signals. Despite the small ($87.42\,\mathrm{MeV}/c^2$) separation between the pole masses of the B^0 and B_s^0 mesons, this separation is much larger than the expected signal resolution and so the two signal peaks can be resolved safely.

The yields from the invariant mass fits are normalised using the result of a fit to a control mode, which, in this analysis, is the decay $B^0 \to J/\psi\, K^{*0}$ where the K^{*0} decays to a $K^+\pi^-$ and the J/ψ to a $\mu^+\mu^-$ pair. The control mode serves several purposes within the analysis with it's primary role being to cancel systematic uncertainties in the ratio of the rare mode and control mode. The decay $B^0 \to J/\psi\, K^{*0}$ is used within this analysis for several reasons. Firstly, it's final state $K^+\pi^-\mu^+\mu^-$ is very similar to the final state of the rare mode, $\pi^+\pi^-\mu^+\mu^-$. This improves the cancellation power of different systematic uncertainties whilst also ensuring that many of the selection requirements used to select the rare mode from the data can also be utilised to select the control mode. Secondly, the decay $B^0 \to J/\psi\, K^{*0}$, is a very pure signal relative to the rare modes having a predicted branching fraction of the order 10^{-3} compared to 10^{-8} for the rare modes. Thirdly, there is a precedent for using $B^0 \to J/\psi\, K^{*0}$ as a control mode within similar rare decay searches such as the confirmed existence of $B_s^0 \to \overline{K}^{*0}\mu^+\mu^-$ [152]. Indeed, due to it's prior use as a control mode, parameters which describe the decay $B^0 \to J/\psi\, K^{*0}$ are well known experimentally. These parameters include the size of the S-wave contribution to the $K^+\pi^-$ final state which describes the proportion of $K^+\pi^-$ daughters which exist in a net zero angular momentum configuration. This particular parameter must be included within this analysis

to align the meaning of the control mode branching fraction used here with that used in previous analyses.

The cancellation power which results from including the control mode normalisation channel can be seen explicitly by first considering that the rare mode yields can be written as

$$\begin{aligned} N(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-) = \mathcal{B}(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-) \\ \times\,\epsilon(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-) \\ \times\, L\times 2\sigma(pp\to b\bar{b})\times f_{d(s)} \end{aligned} \tag{4.1}$$

where $N(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-)$ are the yields, $\mathcal{B}(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-)$ are the branching fractions while $\epsilon(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-)$ are efficiency terms which describe the proportion of candidates which survive the selection procedure. Furthermore, L is the integrated luminosity and $\sigma(pp\to b\bar{b})$ is the cross-section of $b\bar{b}$ production where the additional factor of 2 takes into account the $b\bar{b}$ pair not reconstructed within the LHCb acceptance. Finally, $f_{d(s)}$ is a fragmentation fraction term taking into account the relative production of B^0 and B_s^0 mesons at specific centre of mass energies. By normalising the rare mode yield with a control mode yield the terms describing both the luminosity and cross-section disappear, as seen in Equation 4.2 which is used to extract the final rare mode branching fractions.

$$\begin{aligned} \mathcal{B}(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-) = \frac{N(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-)}{N(B^0\to J/\psi\,K^{*0})}\times\frac{\epsilon(B^0\to J/\psi\,K^{*0})}{\epsilon(B^0(B_s^0)\to\pi^+\pi^-\mu^+\mu^-)} \\ \times\,\frac{\mathcal{B}(B^0\to J/\psi\,K^{*0})\times\mathcal{B}(J/\psi\to\mu^+\mu^-)\times\mathcal{B}(K^{*0}\to K^+\pi^-)}{(1-F_S)} \\ \times\,\frac{f_d}{f_{d(s)}} \end{aligned} \tag{4.2}$$

Within Equation 4.2 the efficiency ratio term, which is determined using Monté Carlo simulated candidates, effectively scales up the measured yields for both the control mode and rare mode to before the application of the analysis selection. In addition, some systematic uncertainties relevant when calculating individual efficiency terms cancel within this ratio. The third term is the branching fraction of the control mode multiplied by the branching fractions of the subsequent decays $J/\psi\to\mu^+\mu^-$ and $K^{*0}\to K^+\pi^-$ and finally multiplied by a term, $(1-F_S)$, where F_S describes the proportion of S-wave $K^+\pi^-$ decays. Finally, the last term of Equation 4.2 denotes the ratio of fragmentation fractions where the subscript d or s refers to a down or strange quark, corresponding to a B^0 or B_s^0 meson. The fragmentation fraction ratio is needed for the

calculation of the B_s^0 mode branching fraction due to the choice of $B^0 \to J/\psi K^{*0}$ as the control mode.

Within the analysis presented within this thesis the value of the control mode branching fraction is taken from [153] and is equal to

$$\mathcal{B}(B^0 \to J/\psi K^{*0}(892)) = (1.19 \pm 0.01 \pm 0.08) \times 10^{-3}, \tag{4.3}$$

where the first uncertainty is statistical and the second systematic. The branching fraction for the J/ψ decay is then taken from [154] and is equal to

$$\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033) \times 10^{-2}, \tag{4.4}$$

while the branching fraction value for the decay $K^{*0} \to K^+\pi^-$ is taken as $2/3$ which is determined from Clebsch-Gordon coefficients. The fraction of S-wave candidates, F_S, is taken from a previous analysis [155] where it was determined to be

$$F_{\rm S} = (6.4 \pm 0.3 \pm 1.0)\%. \tag{4.5}$$

Finally, for the fragmentation fraction values, the ratio f_s/f_d, is measured by LHCb to be 0.259 ± 0.015 for Run 1 [156] and 0.244 ± 0.012 for Run 2 [138]. An alternative value for the Run 2 value can be taken from [157] where the Run 1 value is scaled by 1.068 ± 0.016 to get a different value of 0.277 ± 0.017.

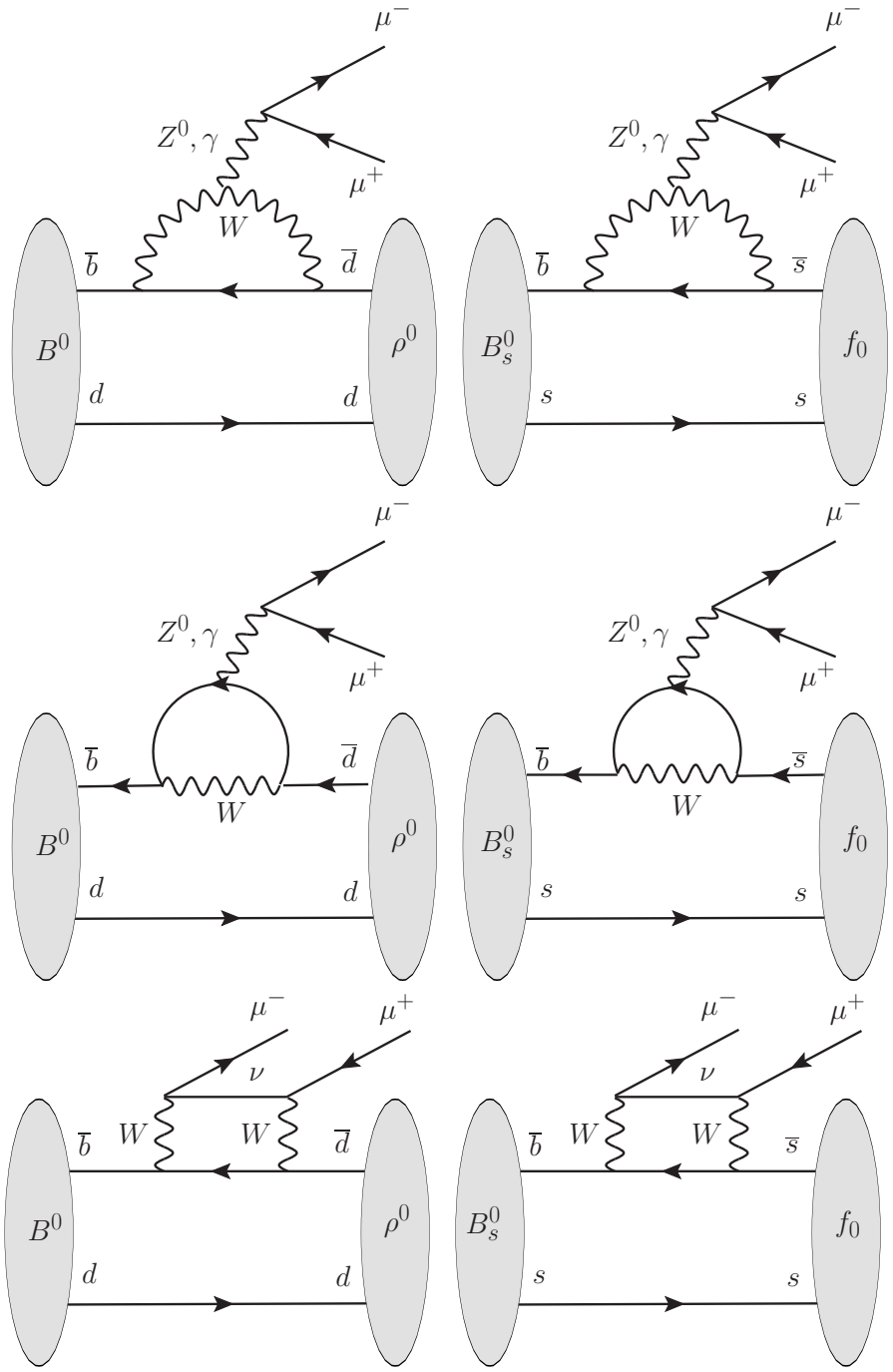

Figure 4.1: Examples of Feynman diagrams for the $b \to d$ process $B^0 \to \rho^0 \mu^+ \mu^-$ (left column) and the $b \to s$ process $B_s^0 \to f_0(980) \mu^+ \mu^-$ (right column). Both final state mesons decay to $\pi^+ \pi^-$. The top and middle row show examples of penguin diagrams while the bottom row shows examples of box diagrams.

4.2 Branching Fraction Variations with q^2 and $m(\pi^+\pi^-)$

In order to make a direct comparison with theoretical predictions, the rare mode branching fractions can be measured in particular regions of q^2 providing sensitivity to different Wilson coefficients. Within this analysis the possibility of performing the branching fraction measurements in the different q^2 regions is explored (in Chapter 5) and separate measurements are made in three different regimes. The first corresponds to the full q^2 range considered, $q^2 \in$ [0.1-19.0] GeV^2/c^4. Within this range candidates within the range $q^2 \in [8.0 - 11.0]\,\mathrm{GeV}^2/c^4$ and $q^2 \in [12.5 - 15.0]\,\mathrm{GeV}^2/c^4$ which are dominated by the J/ψ and $\psi(2S)$ resonances are rejected. The final result in the full q^2 range is corrected to account for the removed regions using a decay model. The second region considered is the low q^2 region which lies below the threshold for the J/ψ charmonium resonance and which corresponds to $q^2 \in$ [0.1-8.0] GeV^2/c^4. The third is the high q^2 region which lies above the $\psi(2S)$ resonance, and which is defined as the q^2 region $q^2 \in$ [15.0-19.0] GeV^2/c^4. In addition, all three of these measurements are made in a bin of the mass of the dipion pair, $m(\pi^+\pi^-)$, which corresponds to $m(\pi^+\pi^-) \in$ [0.515,1.3] GeV/c^2. This choice of mass range is motivated by several factors. Firstly, the number of expected candidates in a higher $m(\pi^+\pi^-)$ region is small making a branching fraction measurement impractical. Secondly, the understanding of the various resonance contributions to the $\pi^+\pi^-$ system at higher values of $m(\pi^+\pi^-)$ is less well known further away from the $\rho^0(770)$ and $f_0(980)$ resonances. This means that any result from this region provides a less useful comparison to theoretical predictions. Thirdly, the lower boundary of the mass bin is set to 0.515 GeV/c^2 to remove backgrounds originating from a K^0_{S} where $K^0_{\mathrm{S}} \to \pi^+\pi^-$. Finally, the mass bin is the same as used by the previous LHCb analysis. This allows a comparison of results within this thesis to be made against previous experimental results.

Candidates which fall within the range $q^2 \in$ [8.0-11.0] GeV^2/c^4, are dominated by decays where the dimuon pair originates from the J/ψ meson. These charmonium decays have branching fractions which are much larger than those attributable to the rare modes. Previous LHCb measurements [158, 159] of these modes have found that the decay $B^0 \to J/\psi\,\pi^+\pi^-$ occurs with a branching fraction of

$$(3.94 \pm 0.14(\mathrm{stat}) \pm 0.21(\mathrm{syst})) \times 10^{-5} \tag{4.6}$$

while $B^0_s \to J/\psi\,\pi^+\pi^-$ has been confirmed with a branching fraction of

$$(2.09 \pm 0.08(\mathrm{stat}) \pm 0.21(\mathrm{syst})) \times 10^{-4}. \tag{4.7}$$

These decays, which are hereafter referred to as the "resonant modes", are orders of magnitude

more common than their rare mode equivalents and yet share a very similar structure. This means their response to the selection process is similar to the rare modes. Furthermore, the signal shapes used within any invariant mass fit to the resonant mode candidates should intuitively be the same as to those of the rare mode candidates. Indeed, many of the experimental backgrounds relevant for a measurement of the yield of the resonant modes, are also valid for the rare mode. These similarities between resonant and rare modes make the resonant modes a very useful proxy and so, prior to the measurement of the rare mode branching fractions, the selection is applied and the branching fractions calculated for the resonant modes. This serves to validate the analysis strategy prior to the measurements of the rare mode.

When confirming the existence of the rare $b \to d$ and $b \to s$ processes it is necessary to demonstrate that the $\pi^+\pi^-$ system originates from either the $\rho^0(770)$ in the case of the B^0 decay, or the $f_0(980)$ in the case of the B_s^0 decay. This is not trivial as there are contributions from various resonances to the $\pi^+\pi^-$ system. Furthermore, due to the large widths of both the ρ^0 (770) and $f_0(980)$ resonances, a simple selection around these resonances in $m(\pi^+\pi^-)$ is not feasible. In order to achieve a confirmation that the dipion system originates from the $\rho^0(770)$ and $f_0(980)$ resonances several techniques could be used. One such technique involves checking the angular distribution of the dipion final state to ensure that the angular information shows consistency with the $\rho^0(770)$ or $f_0(980)$ quantum numbers. Figure 4.2 shows a cartoon showing how the pion helicity angle distribution, $\cos(\theta_h)$, varies with the angular momentum of the intermediate resonance. Within the sketch the red distribution shows the expected distribution if the dipion system originates from an S-wave resonance. Then, the blue and green distributions correspond to P-wave and D-wave resonances which correspond to $J = 1$ and $J = 2$ states respectively. It is therefore theoretically possible to exploit the angular information of the rare mode $\pi^+\pi^-$ system to determine whether the quantum numbers are consistent with the expected rare decay channel.

The existence of the other resonance contributions also has an effect on the accuracy of the estimated efficiencies which determine the final branching fraction values. The Monté Carlo samples which are used to estimate these efficiencies, by default, only consider a single intermediate hadronic resonance component. This means that using the pure LHCb simulation samples for the calculation of these efficiencies will result in an inaccurate final measurement as they do not include the smaller resonant contributions. Within this analysis, this problem is addressed using a weighting procedure where the simulation samples are adjusted, using a model, to artificially include the contributions from the other resonances. The decay model for the resonant mode decays is based upon previous experimental results while the model for the rare mode decays is built following theoretical modelling found in previous literature. Both decay models use an underlying angular basis and the decay models are discussed in the dedicated chapter, Chapter 6.

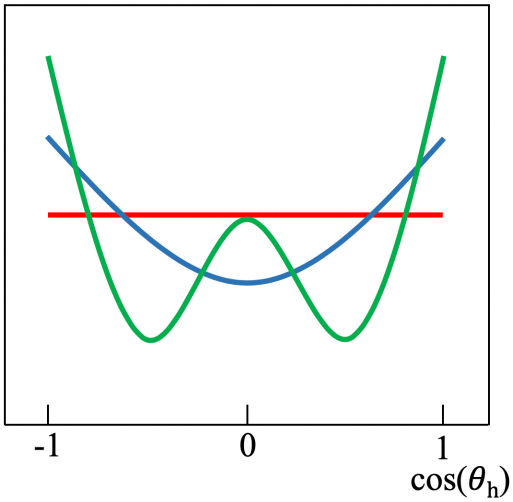

Figure 4.2: A sketch showing the distribution of the pion helicity angle, $\cos(\theta_h)$ for resonances with different quantum numbers. The red distribution shows the generic shape expected for a $J = 0$ state, while the blue and green distributions show what is expected for a $J = 1$ and $J = 2$ state respectively. The sketch is created using the decay model described in Chapter 6.

The remaining chapters of this thesis examine, in detail, the various aspects of the branching fraction analysis which have been introduced here; Chapter 5 describes the selection criteria used within the analysis; Chapter 6 discusses the decay model used for both the resonant and rare modes; Chapter 7 details the methods used to calculate the efficiencies of relevance and Chapter 8 presents the results of the mass fits to the control mode, the resonant mode and the rare mode. The various sources of systematic uncertainty that affect the branching fraction results are detailed in Chapter 9 and, finally, the results and a summary are presented in Chapter 10.

CHAPTER 5

Selection Procedure

“ If you torture the data long enough, it will confess. ”

Ronald Coase,

The branching fraction values of the rare modes $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ are expected to be of the order 10^{-8} and any signal will be hard to identify within the dataset. Furthermore, any measurement of these small signals is frustrated by the presence of numerous experimental backgrounds which are present within the dataset, some of which lie directly on top of the expected signal region in any fit to the $\pi^+\pi^-\mu^+\mu^-$ invariant mass distribution. These backgrounds must be identified and reduced via the application of a series of selection requirements that exploit the decay topologies, kinematics and LHCb sub-detector responses. While the objective of reducing backgrounds is of paramount importance, it is crucial that the number of signal candidates within the dataset remains large enough post-selection to allow for a statistically significant observation of the rare modes. The following chapter discusses the different steps of the selection procedure. Throughout this chapter when referring to a particle decay, the inclusion of charge conjugate processes is implied.

5.1 Trigger

As was discussed within Chapter 3, all candidates within the analysis must have passed both the hardware (L0) and software (HLT1, HLT2) trigger stages, which reduce the LHCb data output rate whilst simultaneously selecting out events of interest. In order to pass each trigger stage, an

event must pass a set of bespoke "trigger lines" which are implemented within the LHCb software and which are only passed under a specific set of topological or kinematic conditions. The set of trigger lines required to have been passed within a given analysis is tailored to the analysis in question with the different trigger lines used within this analysis shown within Table 5.1 where each line is referred to by its name within the LHCb software.

A classification system is used to determine which of the particle tracks within an event are responsible for causing each individual trigger line to fire. This system includes the categories TOS and TIS which stand for "Trigger On Signal" and "Trigger Independent of Signal" respectively. As the names suggest, a line which is fired under a TOS condition is fired in response to a particle track (or combination of particle tracks) that is reconstructed to form part of the signal candidate. Conversely, a line which is fired under a TIS condition is fired due to a particle track (or a combination of particle tracks) elsewhere in the event, which is not reconstructed to form a signal candidate. It is possible for an event to be selected by multiple trigger lines, some of which trigger under a TOS condition and some under a TIS condition. This is referred to as TISTOS and can be used to help evaluate the efficiency of the trigger and this is discussed further in Chapter 7.

In order to pass a given trigger stage, at least one of the trigger lines of each stage must be TOS. As is seen within the table, this requirement is implemented via the application of a logical OR operator between the different trigger lines of each stage.

As can be seen from Table 5.1, the trigger conditions used for the Run 1 data are different from those used for the Run 2 data. This is driven by the different physical conditions of the Run 2 data taking period, the redefinition of trigger lines between the Run 1 and Run 2 data, and also by the availability of new software trigger lines in the Run 2 data which were not available for Run 1. The L0 trigger lines used, for both Run 1 and Run 2, are `L0MuonDecision` and `L0DiMuonDecision`. The `L0MuonDecision` line requires that a pattern of hits in the muon system is present and is consistent with the existence of a track that has a high transverse momentum, p_{T}. The `L0DiMuonDecision` line is triggered by a dimuon pair within the event which has a large product of muon transverse momenta. Both the L0 lines include a requirement on the detector occupancy as an event with too many tracks can take too long, and be too hard, to reconstruct in the high level trigger. The detector occupancy is represented here, by proxy, via the number of hits recorded in the SPD sub-detector. This number of hits is approximately equivalent to the number of charged particles within the event which travel far enough downstream to reach the SPD. The value of this occupancy requirement is different between the two L0 trigger lines. The requirement for the `L0MuonDecision` trigger line is that the number of SPD hits be < 600 for Run 1 and < 450 for Run 2. The equivalent requirement for the `L0DiMuonDecision` line is that the number of SPD hits be < 900.

Table 5.1: The trigger requirements which candidates are required to have passed for the Run 1 and Run 2 data. For each stage of the trigger (L0, HLT1, HLT2) a candidate must pass at least one of the lines, implemented via a logical OR operator.

Trigger Stage	Run 1	Run 2
L0	L0MuonDecision or L0DiMuonDecision	L0MuonDecision or L0DiMuonDecision
HLT1	Hlt1TrackAllL0Decision or Hlt1TrackMuonDecision	Hlt1TrackMuonDecision or Hlt1DiMuonHighMassDecision or Hlt1DiMuonLowMassDecision or Hlt1TrackMVADecision or Hlt1TwoTrackMVADecision
HLT2	Hlt2Topo2BodyBBDTDecision or Hlt2Topo3BodyBBDTDecision or Hlt2Topo4BodyBBDTDecision or Hlt2TopoMu2BodyBBDTDecision or Hlt2TopoMu3BodyBBDTDecision or Hlt2TopoMu4BodyBBDTDecision or Hlt2DiMuonDetachedDecision	Hlt2DiMuonHighMassDecision or Hlt2SingleMuonDecision or Hlt2SingleMuonRareDecision or Hlt2Topo2BodyDecision or Hlt2Topo3BodyDecision or Hlt2Topo4BodyDecision or Hlt2TopoMu2BodyDecision or Hlt2TopoMu3BodyDecision or Hlt2TopoMu4BodyDecision or Hlt2TopoMuMu2BodyDecision or Hlt2TopoMuMu3BodyDecision or Hlt2TopoMuMu4BodyDecision or Hlt2DiMuonDetachedDecision

The HLT1 trigger lines select out decays which have particles with high p_{T} values, a good track quality and a high impact parameter, (IP) [160]. Lines which specifically target muons, such as `Hlt1TrackMuonDecision` have slightly relaxed thresholds on these requirement values in order to increase the trigger efficiency for decays containing final-state muons. For Run 2 the `TrackMVA` lines which use a Multi-Variate Analysis (MVA) to remove tracks with a low reconstruction quality were added. This MVA takes as inputs the number of hits in the sub-detectors, the reconstruction quality of both individual track segments and the total reconstructed track and the track p_{T} value. The HLT1 lines used within this analysis also include the `DiMuonMass` lines which require a pair of muons and which place requirements on the invariant mass of the dimuon pair and also place requirements on the quality of the dimuon vertex reconstruction.

The HLT2 trigger lines used include multiple `Topo` lines [161]. These trigger on candidates which exhibit a topology commonly shared by heavy b-hadron decays such as the characteristic displaced secondary vertex and the high p_{T} and IP value. The lines make use of a

Boosted Decision Tree (BDT) algorithm to make the final trigger decision by taking into account the detached secondary vertex information, lifetime information, IP values and p_{T} information simultaneously. Other HLT2 lines which are used include the `Hlt2DiMuonDetached` line which selects events containing a dimuon pair that is significantly displaced from the PV and the `Hlt2SingleMuonRare` line which includes requirements to select a high p_{T} and IP single muon track and which is tailored towards rare decay searches.

5.2 Stripping

The huge amount of data accumulated by the LHCb experiment goes through many different stages of processing before it is analysed. Although much of the selection procedure varies from analysis to analysis, a certain stage of the selection process, known as the stripping selection, is more standardised. This is a necessity given the unwieldy size of the entire LHCb dataset. The stripping selection is a series of loose requirements made on kinematic, topological, quality of reconstruction and global event variables. The objective of these requirements is to reduce the dataset to a manageable size for off-line analysis. In addition the stripping provides a common starting point for physics analyses which will share certain immutable characteristics, such as, for example, the presence of a dimuon pair within the final-state. A set of stripping requirements is known collectively as a stripping "line" and the data which pass similar stripping lines are grouped together into data "streams". It is therefore possible to visualise the full LHCb dataset as being divided up into different data streams which broadly describe very general characteristics of a decay before the streams are subdivided further via the relevant stripping lines into a smaller dataset more tailored for a specific physics search.

The analysis presented within this thesis makes use of the LHCb `Leptonic` stream, which selects events with at least one lepton within the final-state. More specifically, the analysis makes use of the LHCb `B2XMuMu` stripping line, which looks for events containing a b-hadron that decays to an oppositely charged dimuon pair within the final-state and a hadron, denoted by X. This stripping line is widely used for $b \to s(d)\mu^+\mu^-$ decay searches.

The selection requirements made by the `B2XMuMu` stripping line can be found in Table 5.2. Some of the requirements within this table differ from year to year as LHCb stripping requirements are decided upon by the collaboration on an annual or biannual basis as part of dedicated "stripping campaigns". The `B2XMuMu` line places a series of requirements on the invariant mass of the reconstructed B^0 (B_s^0) meson and also on the χ^2 of the distance of closest approach (DOCA) of the B meson trajectory to the PV, the χ^2 flight distance between the B candidate and associated PV and quality of vertex χ^2. A requirement is also placed on the size of the DIRA angle, which is defined as the angle between the momentum vector of the reconstructed particle and the vector

Table 5.2: Stripping requirements applied by the LHCb `B2XMuMu` stripping line as part of the `Leptonic` stream. Where the requirements vary depending on the year of data taking, the relevant year is specified in brackets. $m(\pi^+\pi^-\mu^+\mu^-)$ and $m(\pi^+\pi^-)$ refer to the invariant mass combinations of the $\pi^+\pi^-\mu^+\mu^-$ and $\pi^+\pi^-$ systems respectively.

Particle	Variable	Cut Value
B^0	IP χ^2	< 16
	$m(\pi^+\pi^-\mu^+\mu^-)$	$> 4900\,\text{MeV}$ (11,12,15,17,18)
	$m(\pi^+\pi^-\mu^+\mu^-)$	$> 4700\,\text{MeV}$ (16)
	$m(\pi^+\pi^-\mu^+\mu^-)$	$< 7100\,\text{MeV}$ (11,12,15,16)
	$m(\pi^+\pi^-\mu^+\mu^-)$	$< 7000\,\text{MeV}$ (17,18)
	cos(DIRA angle)	< 0.9999
	Flight Distance χ^2	> 121
	Vertex χ^2/ndf	< 8
$\pi^+\pi^-$	$m(\pi^+\pi^-)$	$< 6200\,\text{MeV}$
	DOCA χ^2	< 20
	Vertex χ^2/ndf	< 12
	DIRA	$>$ -0.9
	Flight distance χ^2	> 9 (11,12) 16 (15,16,17,18)
$\mu^+\mu^-$	$m(\mu^+\mu^-)$	$< 7100\,\text{MeV}$
	Vertex χ^2/ndf	< 12
	DIRA	$>$ -0.9
	Flight distance χ^2	9
$\mu^\pm$	`IsMuon`	True
	$\text{DLL}_{\mu\pi}$	$>$ -3
All Tracks	Ghost Probability	< 0.35 (11,12) 0.5 (15,16,17,18)
	Min. IP χ^2 w.r.t all PVs	> 9
Global Event Cut	SPD multiplicity	< 600

from the particle's origin vertex to its decay vertex. Figure 5.1 shows the IP, flight distance and DIRA angle in the context of the decay topology of a $B^0(B_s^0) \to X\mu^+\mu^-$ process. Within the schematic a generic intermediate resonance is represented by X and generic final-state hadron by $h^\pm$ which represents either pion or kaon. The four momentum vector shown belongs to the b-hadron and is used as an example.

The stripping requirements place additional requirements on the intermediate vector meson and on the dimuon pair and particle identification requirements are made on the two final-state muons. These include a requirement on the likelihood that each muon track is a muon with respect to the likelihood that it is a pion, described by the difference in log-likelihoods variable, $\text{DLL}_{\mu\pi}$ which provides muon-hadron separation. The identification requirements also

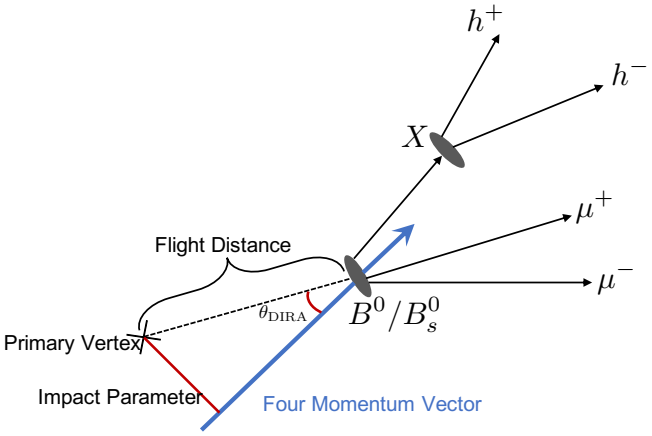

Figure 5.1: The topology of the rare mode decays. The impact parameter (IP), flight distance (FD) and DIRA angle, $\theta_{\rm DIRA}$ are labelled. The labels X and $h^{\pm}$ refer to generic intermediate hadronic resonances and final-state hadrons (kaons or pions) respectively. For a genuine signal, $\theta_{\rm DIRA} = 0$.

include the need for the final-state muons to satisfy the `IsMuon` requirement which is a low-level muon PID variable created by checking whether the track can be associated with hits in any of the LHCb muon chambers [162]. The `B2XMuMu` stripping line also places requirements on the probability that each track is in fact a "ghost" track which refers to a track which is reconstructed from the partial trajectories of more than one particle. This probability is the output of a neural network algorithm which takes into consideration vertex quality, IP and $p_{\rm T}$ information [163]. A general requirement is then placed on the quality of all decay vertices present within the event and, finally, on the number of hits recorded within the SPD sub-detector. This occupancy requirement is aligned with the `L0Muon` occupancy requirement in Run 1.

5.2.1 Post-stripping

After the application of the stripping selection it is possible to observe some rudimentary structures within the data from the invariant $\pi^+\pi^-\mu^+\mu^-$ mass distributions. This is a useful exercise because even at this early stage of the selection, the profile of many of the dominant backgrounds are visible. Figure 5.2 shows the invariant mass distributions, post-stripping. The top plots show the invariant mass distribution when selecting only candidates that are attributable to the resonant modes. The bottom plots show the mass distributions when vetoing these candidates and selecting for the rare modes. Both of these types of distribution show a broad exponential shape which decreases in it's contribution as the invariant mass is increased. This shape can be attributed to combinatorial background which is a form of background that results from the reconstruction of

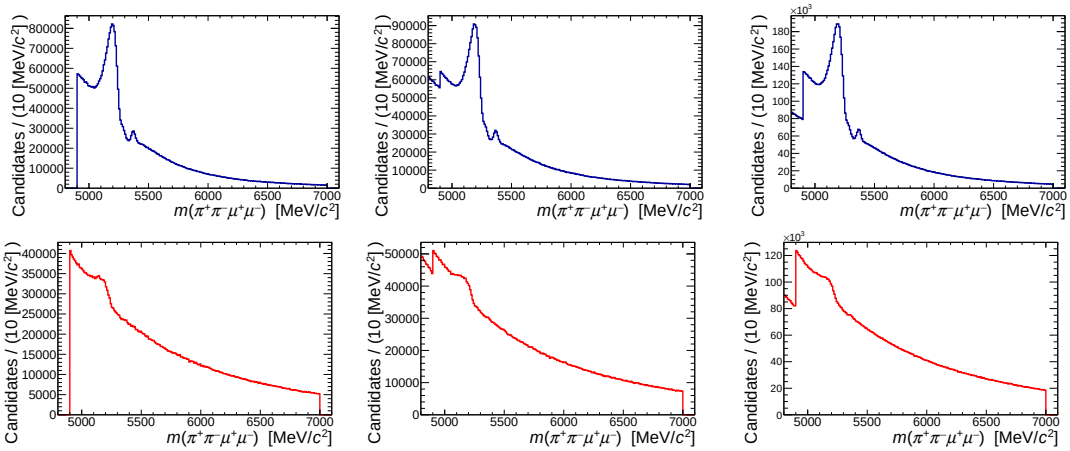

Figure 5.2: The invariant $\pi^+\pi^-\mu^+\mu^-$ mass distributions after the application of the stripping selection, for the 2011+2012 dataset (left column), 2015+2016 dataset (middle column) and 2017+2018 dataset (right column). The top plots show the reconstructed candidates where the dimuon pair originates from the J/ψ resonance and the lower plots the equivalent where only candidates where the dimuon pair does not originate from the J/ψ resonance are included.

random combinations of particle tracks within an event that manage to pass the initial selection criteria. Given it's prevalence the selection procedure must reduce this form of background as much as possible and this is addressed in Section 5.4. The plots from Figure 5.2 attributable to the Run 2 datasets show the existence of discontinuities within the mass spectrum. These are present due to the different $m(\pi^+\pi^-\mu^+\mu^-)$ stripping requirements made on data in different years, as seen in Table 5.2.

The resonant mode distributions from Figure 5.2 show two peaks which can be observed above the broad combinatorial exponential. The larger peak is dominated by the decay mode $B^0 \rightarrow J/\psi K^{*0}$ and is present within the dataset due to the inclusion of events where the selection criteria mis-identify a kaon from the event as a pion before going on to reconstruct a $K^+\pi^-\mu^+\mu^-$ candidate. The incorrect identification of the kaon as a pion results in this peak being smeared out and then shifted within the distribution to lower values than the known B^0 mass of $5281\,\mathrm{MeV}/c^2$ [154], due to the difference in kaon and pion masses. As can be seen, this is a dominant source of background within the analysis whose importance is driven by its large branching fraction compared to the corresponding decays to a $\pi^+\pi^-$ final-state (indeed, this purity motivates it's selection as a control mode for the analysis). This source of background is reduced using particle identification requirements discussed in Section 5.5. However, as the background lies in the same invariant mass region as the B^0 signal mode it is modelled as part of the final invariant mass fits. In addition to this large peak, a smaller peak can be seen to the right,

around 5360 MeV/c^2, which is attributable to the $B^0_s \to J/\psi\,\pi^+\pi^-$ decay. The $B^0 \to J/\psi\,\pi^+\pi^-$ decay, is not visible without vetoing the $B^0 \to J/\psi\,K^{*0}$ background.

The lower plot of Figure 5.2 shows far fewer structures but is once again obscured by combinatorial background. A wide structure can be seen around 5200 MeV/c^2, which is attributable to the decay $B^0 \to K^{*0}\mu^+\mu^-$, entirely analogously to the $B^0 \to J/\psi\,K^{*0}$ background within the resonant mode distribution. The B^0 and B^0_s signal from the two rare modes, prior to the application of the full selection, are obscured.

At this stage of the analysis a requirement is placed on the opening angles between the two positive tracks which are included within the candidate and the equivalent angle between the two negative tracks. The requirement rejects candidates where these angles are smaller in size than 0.5 mrad. The purpose of these requirements is to reject events which include "clone tracks" where an individual track is used twice to reconstruct a candidate.

5.3 Monté Carlo Simulation

Throughout the analysis, data samples originating from a Monté Carlo (MC) simulation are used in order to optimise each step of the selection process, provide an estimation of efficiencies, train the BDT used to separate signal from background and determine the shapes which describe the signal and background components in final invariant mass fits.

Dedicated MC samples are generated using the full LHCb simulation software (with the software itself described in slightly more detail in Appendix A). The generated MC samples model the decays $B^0 \to \rho^0\mu^+\mu^-$, $B^0_s \to f_0(980)\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$, $B^0 \to J/\psi\,K^{*0}$, $B^0 \to J/\psi\,K^+\pi^-$, $B^0_s \to J/\psi\,f_0(980)$, $B^0 \to J/\psi\,\rho^0$, $B^0_s \to J/\psi\,\eta'$ and $B^0_s \to J/\psi\,\phi$. These MC samples, apart from the $B^0_s \to J/\psi\,\eta'$ and $B^0_s \to J/\psi\,\phi$ samples which are background modes where a final-state particle is not reconstructed, share the same topology as shown in Figure 5.1. These two background modes are relevant to the resonant mode fits of Chapter 8. Due to the large number of potential backgrounds of potential relevance to both the resonant and rare modes, additional simulation samples, such as for the decay mode $B^0_s \to \eta'\mu^+\mu^-$, are also useful to this analysis. However, due to limited computing resources these additional simulation samples are not generated using the complete LHCb MC software. Instead, they are generated with the `RapidSim` software package [164] which is a particle-gun generator which simulates only the decaying particle instead of the underlying event. `RapidSim` determines the kinematic distributions of the generated particles by sampling histograms generated from fixed-order next-to-leading-logarithm transverse momentum calculations [165]. `RapidSim` then decays the generated particles inside a bespoke geometry using physics models from the `EvtGen` package. All MC samples used within the analysis are initially truth-matched whereby the information

from the generator level samples, which is the MC sample prior to passing through the full LHCb data-flow, is compared to the final output after passing through the LHCb digitisation, reconstruction and stripping stages. Candidates that are included within the simulated datasets as a result of an artifact of the generation, instead of the decay of interest, are then removed before the simulation is used for analysis.

The MC samples for the $B^0 \to \rho^0\mu^+\mu^-$, $B^0_s \to f_0(980)\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ are referred to throughout this thesis as the $B^0 \to \pi^+\pi^-\mu^+\mu^-$, $B^0_s \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ samples respectively. This is because these LHCb MC samples are weighted within Chapter 6 to include intermediate resonances which contribute to the $\pi^+\pi^-$ ($K^+\pi^-$) system in addition to the ρ^0 (770), $f_0(980)$ and K^{*0} (892) which they are generated with by the LHCb collaboration. This is also the case for the resonant mode MC samples $B^0_s \to J/\psi\, f_0(980)$ and $B^0 \to J/\psi\, \rho^0$ which are referred to as $B^0_s \to J/\psi\, \pi^+\pi^-$ and $B^0 \to J/\psi\, \pi^-\pi^-$ after applying the decay model weights.

5.3.1 Data/Simulation Agreement and Weighting

Given the importance placed on the MC simulation data samples within the analysis it is imperative that these simulated samples model the data accurately. Although this agreement is generally good across the majority of variables of interest to the analysis there are specific variables that are known to show a poorer level of agreement as they are not properly reproduced by the `PYTHIA` package and the LHCb simulation. These include the $p_{\rm T}$ of the B^0 (B^0_s), the vertex χ^2 of the B^0 (B^0_s) and the detector occupancy, which describes the number of tracks or hits registered by the detector.

In order to address this data-MC disagreement, the MC samples used within the analysis are corrected via a sequential weighting of the B^0 $p_{\rm T}$, vertex χ^2 fit and event occupancy variables. This procedure first requires the separation of signal from background within the data distributions. This separation is achieved using the *sPlot* technique [166] where weights are extracted from a fit to the $B^0 \to J/\psi\, K^{*0}$ control mode invariant mass distribution. The set of *sPlot* weights are then applied to the data to extract the "true" variable distributions and these distributions are then compared to the MC distributions to calculate the final set of kinematic weights. The application of these weights on the MC distributions is seen to dramatically improve the data-MC agreement as can be seen in Figure 5.3 where the un-weighted MC distributions are shown in red, the unfolded data distributions in black and the weighted MC distributions in blue.

The *sPlot* technique exploits the fact that by clearly separating signal from background within one variable distribution, the signal and background distributions can be separated from one another in a second variable assuming that the covariance relation between the two variables

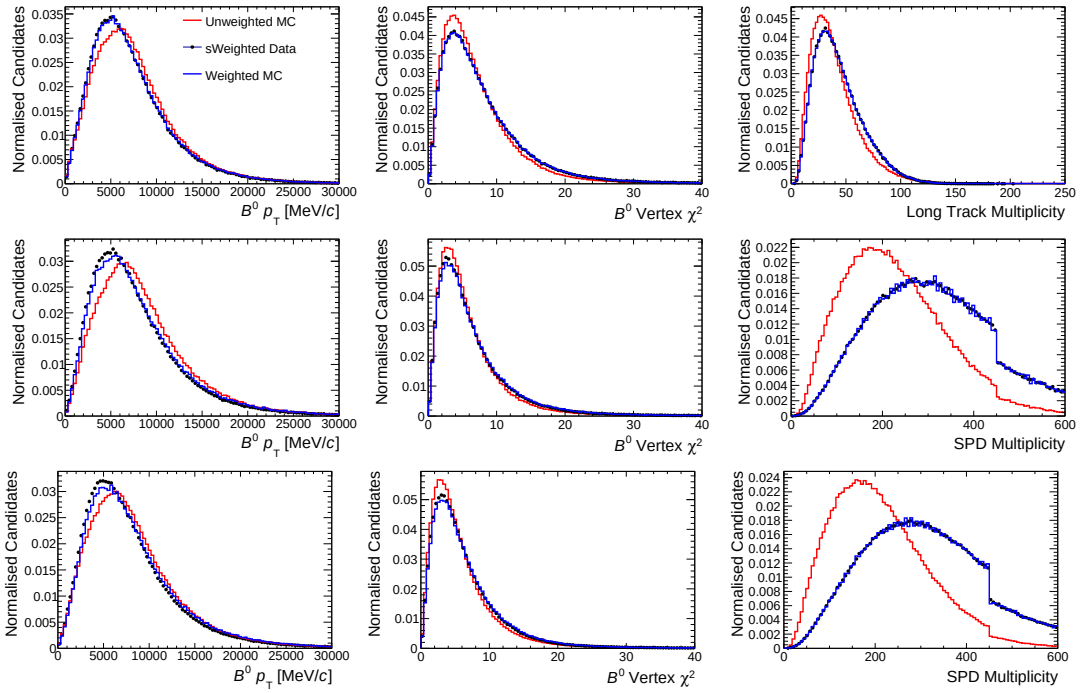

Figure 5.3: The result of the MC kinematic variable weighting for the 2011+2012 dataset (top), the 2015+2016 dataset (middle) and the 2017+2018 dataset (bottom) in the B^0 p_{T}, B^0 vertex χ^2 and event occupancy variables. A good level of agreement can be seen between the unfolded data distributions and the weighted $B^0 \to J/\psi\, K^{*0}$ MC distributions.

can be determined [166]. In practice, the *sPlot* weights take both positive and negative values to "up-weight" candidates more likely to belong to the signal distribution while "down-weighting" candidates more likely to belong to the background distribution. When applying the weights to the second variable the effect is to artificially remove the candidates most likely to belong to background. Within this analysis the first variable is the invariant $m(K^+\pi^-\mu^+\mu^-)$ mass from the control mode, while the second variables are the three kinematic variables to be corrected. Crucially, this technique is only applicable in this context as the invariant mass is not highly correlated with the three kinematic variables.

Different variables are used between Run 1 and Run 2 to describe the detector occupancy. In Run 1 this variable is the number of long tracks in the event, whereas in Run 2, due to changes in track reconstruction, the number of hits in the SPD sub-detector is used as it is thought to be a more reliable proxy. Another feature seen in the figure is that there is a discontinuity which appears in the Run 2 occupancy variable distributions. As was previously mentioned within Section 5.1 there exists a difference between the occupancy requirements included as part of the two L0 trigger lines used, `L0MuonDecision` and `L0DiMuonDecision`, in Run 2. This

difference manifests itself as a discontinuity in the distribution where any candidates with a number of SPD hits above 450 are only passing the `L0DiMuon` trigger requirement.

In addition to this weighting procedure, additional corrections are made to the MC samples to further improve the agreement between simulation and data. Corrections are applied to correct the MC particle identification efficiencies, discussed in Section 5.5, and to correct the calculated trigger and tracking efficiencies as is discussed in Chapter 7. Furthermore, corrections are applied to improve the accuracy of the underlying MC physics models and this is discussed in Chapter 6.

5.4 Multi-Variate Analysis

The large levels of combinatorial background seen within the data cannot be isolated using standard requirements or vetoes as would be possible for other background sources. This is due to the fact that the combinatorial background and signal show very similar behaviour across many physical variables meaning it is not possible to separate the two using a single variable. In order to isolate and reduce this background it is necessary to make use of Multi-Variate Analysis (MVA) techniques that take in many different variables into account simultaneously order to make an informed decision on whether an individual candidate is likely, or not, to be attributable to signal.

Many tools exist which can be used for a multivariate analysis due to recent renewed interest in machine learning and artificial intelligence. Within this analysis, the reduction of combinatorial background is performed using a gradient-Boosted Decision Tree (BDT) [167] classifier implemented via the XGBoost algorithm [168].

5.4.1 XGBoost: Training

In order to train the XGBoost classifier to separate signal and background events the classifier must first be trained to recognise the different patterns within the data which separate out these two categories. The training is performed with two proxy samples, one of which contains examples of signal candidates and the other containing examples of background candidates. The MC sample of $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decays is used as the signal proxy. The background sample is taken from the upper sideband of the data, which corresponds to candidates with an invariant $\pi^+\pi^-\mu^+\mu^-$ mass of greater than $5800\,\mathrm{MeV}/c^2$. This upper sideband region is not expected to contain any significant signal or peaking background features and is therefore dominated by combinatorial background. For this reason candidates in the upper mass sideband are not used in the subsequent analysis. The two proxy samples are added together to form a larger sized sample and then 75% of this larger sample is used to train the classifier. The remaining 25% is used to test the

performance of the classifier and to ensure that the levels of over-training present are small.

Table 5.3: Variables used to train the XGBoost BDT.

Particle	Variable	BDT Ranking Run 1	BDT Ranking Run 2
B^0 (B_s^0)	Vertex χ^2	5	2
	cos(DIRA angle)	10	9
	IP χ^2	12	8
	Isolation	11	12
	Lifetime	4	5
	p_{T}	1	1
$\mu^+\mu^-$	IP χ^2	9	10
$\pi^+\pi^-$	Minimum IP χ^2	2	6
$\mu^\pm$	Minimum IP χ^2	6	3
	Minimum p_{T}	3	4
	$\lvert\mu^+ p_{\mathrm{T}} - \mu^- p_{\mathrm{T}}\rvert$	8	11
Daughters	Minimum p_{T}	7	7

The classifier is trained using a set of training variables which are listed in Table 5.3. These variables are selected for the training procedure due to the fact that they show different behaviour between the signal and background proxy samples, as can be seen in Figure 5.4 and Figure 5.5. Due to differences in these variables between the Run 1 and Run 2 data samples, which exist due to differences in the physical data-taking conditions, the XGBoost training is completed twice: once for the Run 1 data sample and once for the sample corresponding to the data taking years 2015 and 2016. Due to similar variable behaviour seen across Run 2, the 2017+2018 data has the 2015+2016 classifier applied to it.

One of the training variables included, which has not yet been defined, is the "Isolation" training variable. This variable considers tracks, not reconstructed as part of the signal candidate, which fall in a cone around the signal candidate. For each non-signal track in the cone the signal candidate is refit with one of the signal tracks replaced with one of the non-signal tracks. The difference in vertex χ^2 is calculated between the nominal fit and the fit with track replacement. Then, after all possible replacements are made the minimum difference in χ^2 seen is converted to a probability. The separation power of this variable comes from the fact that, for a signal-like event, the signal candidate should be relatively isolated from non-signal tracks. Such a candidate would register a lower value of the Isolation variable than a background candidate formed by particles from two different b-hadron decays.

The final two columns within Table 5.3 show the ranking of each variable which describes how "useful" the algorithm finds each variable when separating signal from background. Within

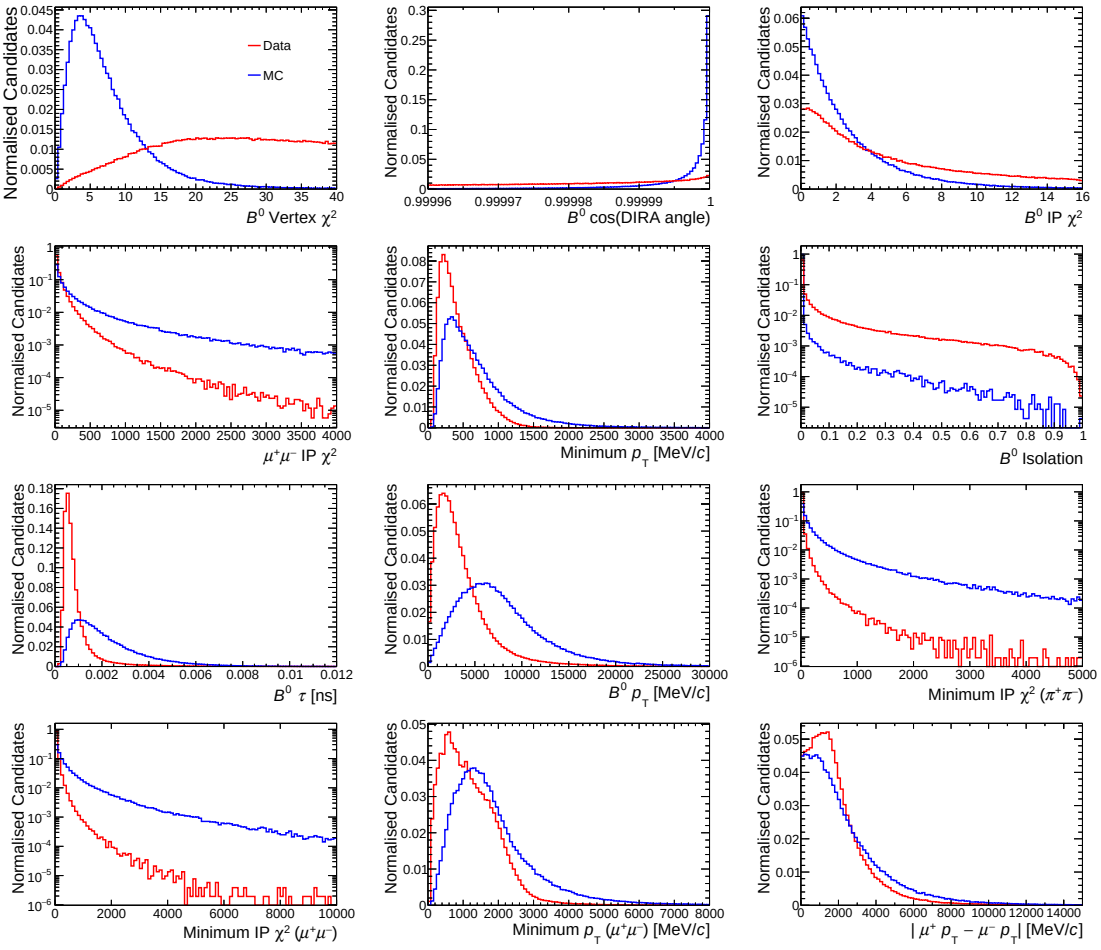

Figure 5.4: The input variables used for the BDT training for Run 1. The variables are shown for the $B^0 \to J/\psi K^{*0}$ mode. Blue shows the signal distribution represented by MC and the red shows the background taken from the upper mass sideband of the data. All histograms are normalised to unity.

this context the usefulness is determined by counting how many times each variable was cut on within the training procedure. For example, if a single decision tree made two requirements on the value of the Isolation variable and one requirement on the value of the mother p_{T}, then the ranking of the Isolation variable is higher than that for the transverse momentum. This ordering is then summed across all of the training trees to produce a final ranking.

Given that the background proxy sample is selected from the data by making a requirement on the invariant $\pi^+\pi^-\mu^+\mu^-$ mass, it is important to ensure that the BDT cannot "cheat" by deciding that all background events can be simply categorised by making a requirement on this

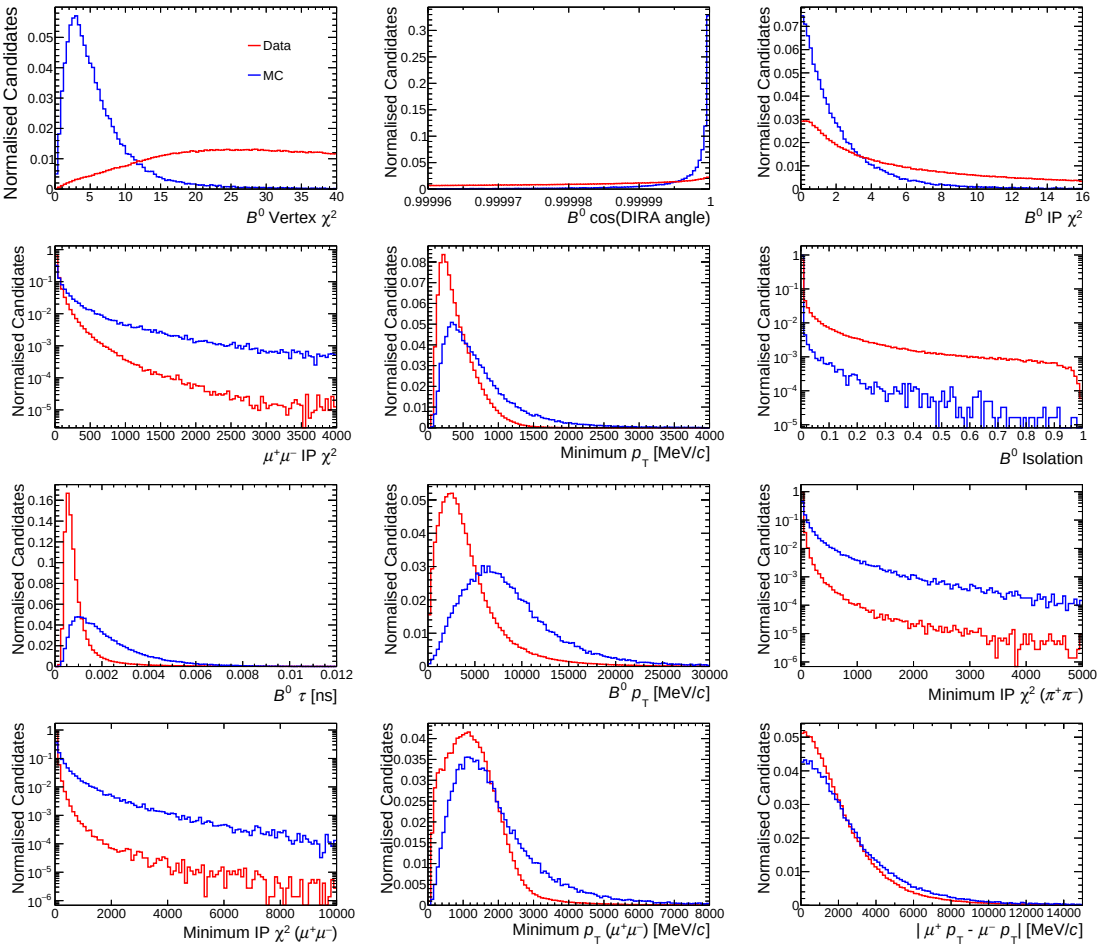

Figure 5.5: The input variables used for the BDT training for 2015+2016. The variables are shown for the $B^0 \to J/\psi K^{*0}$ mode. Blue shows the signal distribution represented by MC and the red shows the background taken from the upper mass sideband of the data. All histograms are normalised to unity.

mass variable. Although this variable is not included within the training directly, there is a danger that the BDT could reconstruct it if it is highly correlated with one or more of the training variables. To ensure that this is not the case, 2-D histograms are filled with the training variable distributions and the invariant mass distribution. These histograms which can be seen for the Run 1 dataset in Figure 5.6, do not show any significant correlations and the same lack of correlation is seen for the Run 2 dataset.

The accuracy of the final BDT classification decision relies on the quality of the training samples which are used to train the algorithm. This includes the relative size of the samples, the

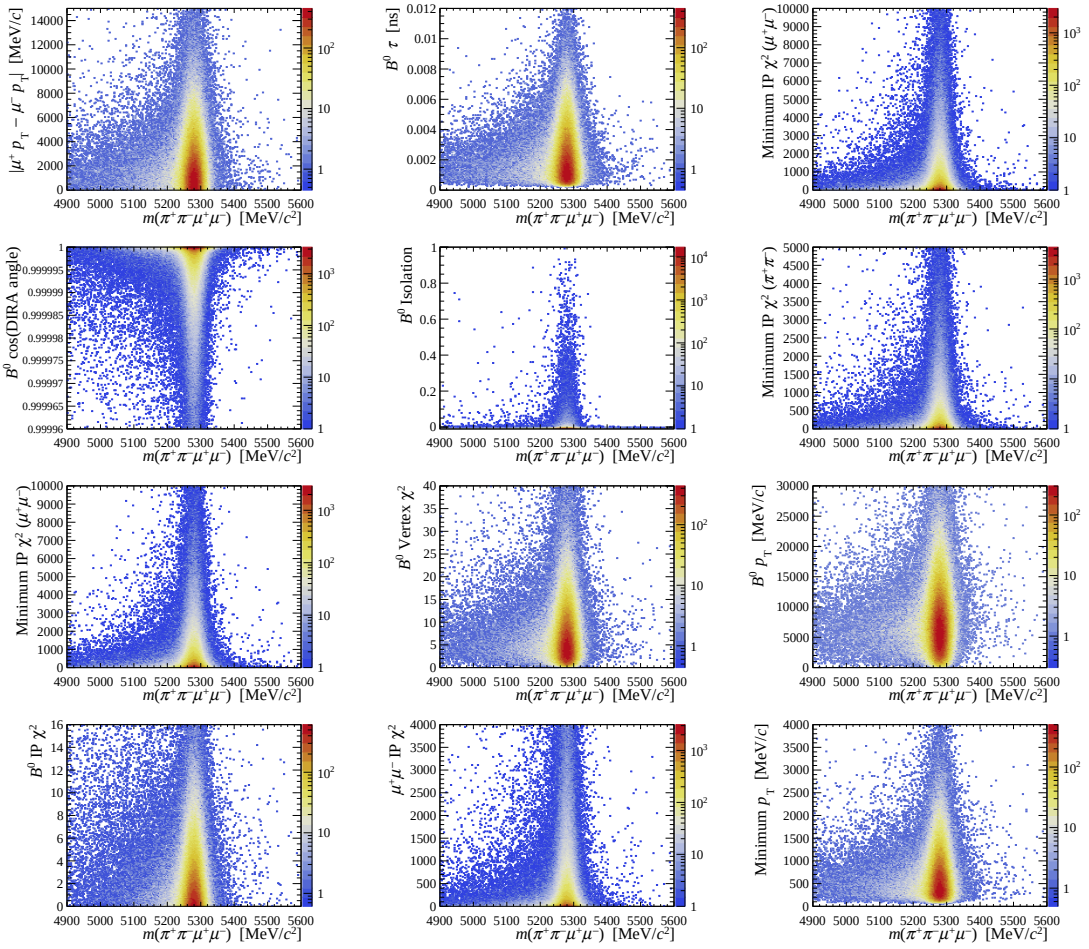

Figure 5.6: The correlations of the BDT training variables with the $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass taken from the Run 1 dataset. The lack of clear correlation indicates that the variables can be safely used within the training of the classifier. Similar plots are seen for the Run 2 datasets.

separation power of the variables and the ability of the MC data samples to accurately reflect the behaviour seen in the data. As was discussed in Section 5.3, the MC samples used within the analysis are weighted in different kinematic variables to improve the data-MC agreement. The level of data-MC agreement for the training variables used is checked, both before and after this weighting, to ensure that the variables reliably match the data and these distributions can be seen in Figure 5.7 and Figure 5.8. The agreement between MC and data is seen to be very good for the training variables however small levels of disagreement can be seen in the distributions which rely on the transverse momentum of the final-state particles. This is understood to be due to the model which is used to generate the $B^0 \to J/\psi\, K^{*0}$ MC sample used to calculate the *sPlot*

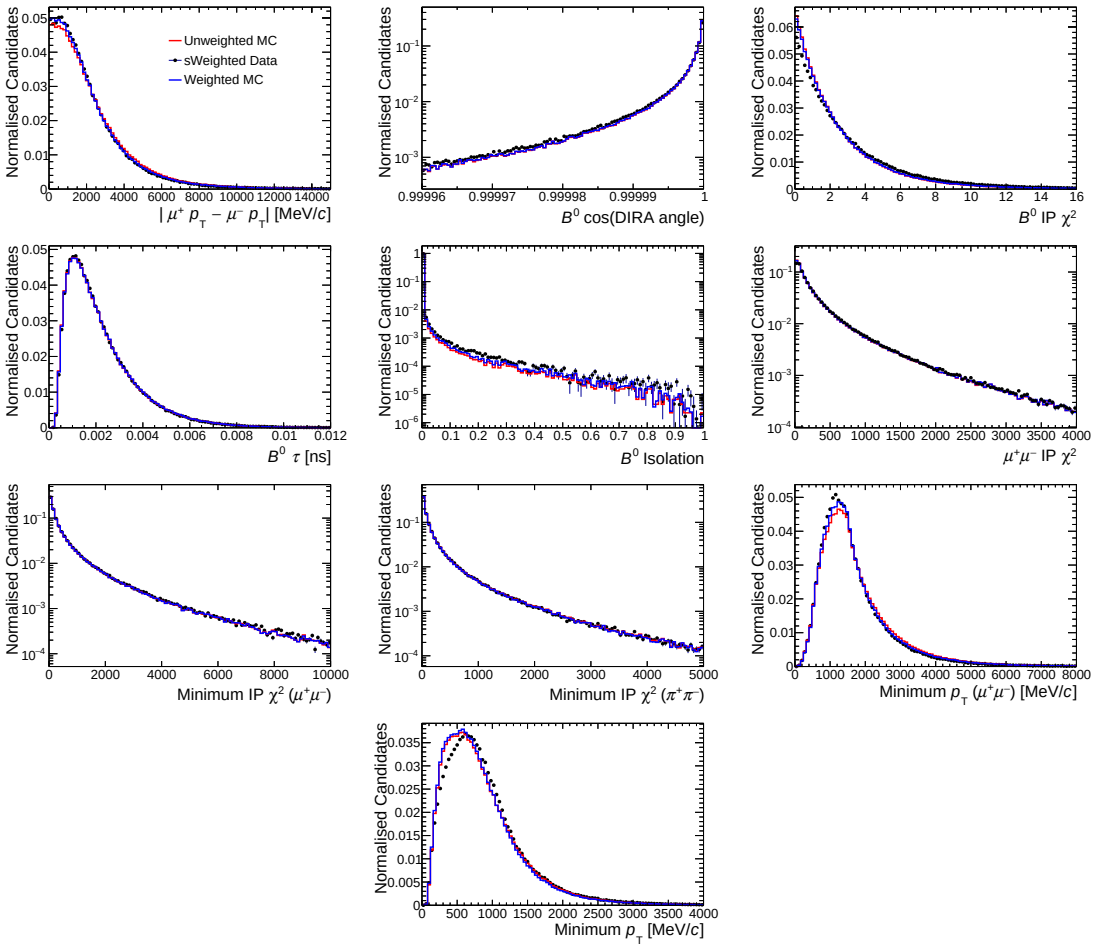

Figure 5.7: The data-MC agreement seen for the Run 1 BDT training variables. The original MC sample distributions are shown in red and the MC sample after the kinematic weighting in blue. The data which is unfolded using the *sPlot* method to make the data-MC comparison is represented by the black data points. For all of the plots shown the area of the histograms is normalised to unity.

weights. The model used generates decays via a P-wave $K^{*0}(892)$ resonance. In reality, is is known that a fraction of these decays proceed in an S-wave configuration and this discrepancy accounts for the limited level of disagreement still seen after the kinematic weighting. Indeed, this discrepancy can be seen to be removed after the application of the decay model weights from Chapter 6, as can be seen in Figure 5.9. Ultimately, the effect of this discrepancy during the training makes very little difference on the total BDT classifier as it is taken into consideration with all of the other variables together.

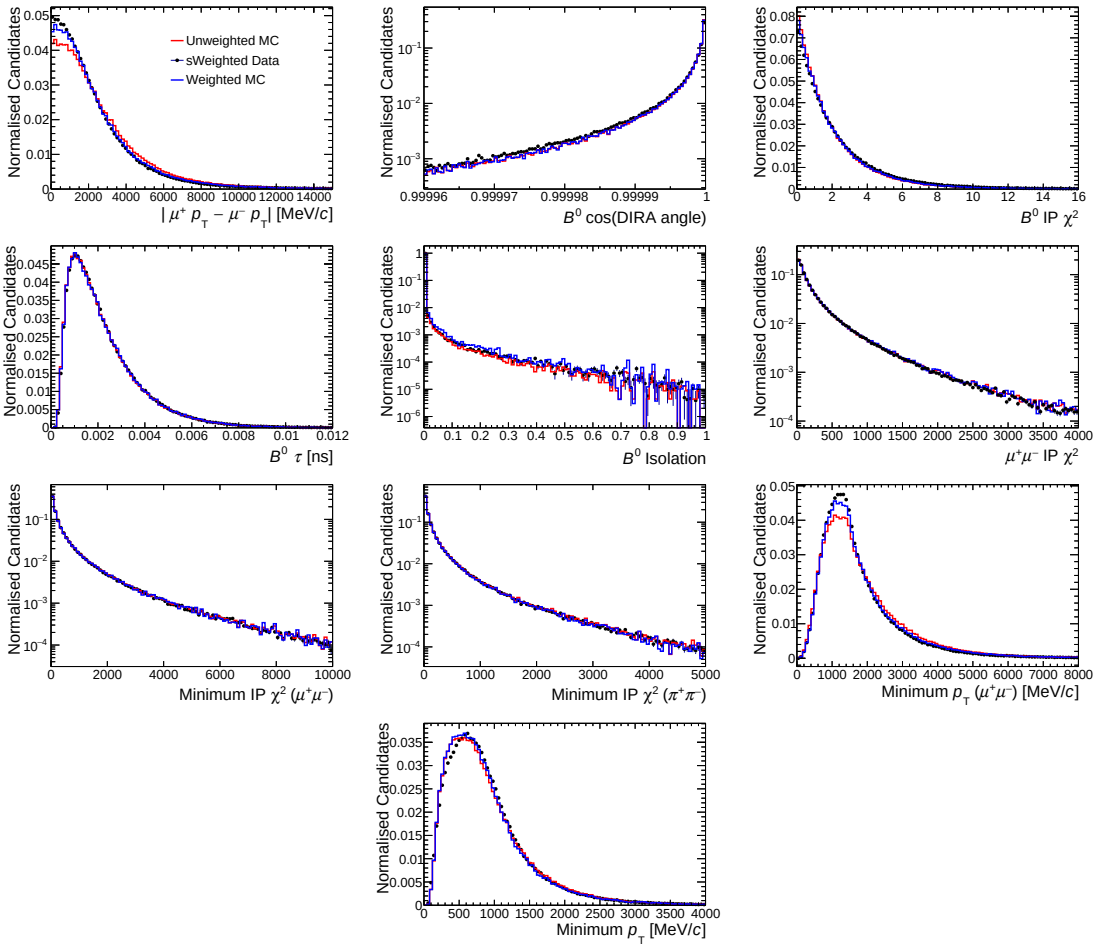

Figure 5.8: The data-MC agreement seen for the 2015+2016 BDT training variables. The original MC sample distributions are shown in red and the MC sample after the kinematic weighting in blue. The data which is unfolded using the *sPlot* method to make the data-MC comparison is represented by the black data points. For all of the plots shown the area of the histograms is normalised to unity.

In addition to the training variables used within the BDT, additional variables, known as "hyper-parameters", exist which define the structure of the BDT classifier itself. Three different hyper-parameters are used within the training of the XGBoost classifier and these are the maximum depth of the training trees, the number of rounds of training performed and the classifier learning rate. These parameters can be adjusted to improve the performance of the training while simultaneously avoiding large amounts of bias and over-training. The maximum depth of the individual training trees was chosen to be 4, the number of rounds of training to be

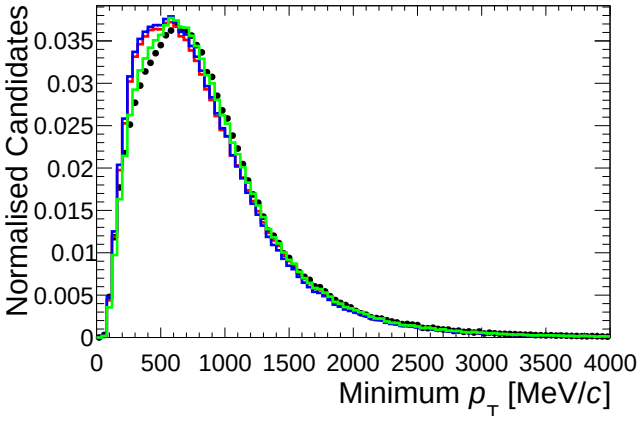

Figure 5.9: Improvement seen in the data-MC agreement in the minimum p_{T} variable when including the decay model weights of Chapter 6. The data which is unfolded using the *sPlot* method to make the data-MC comparison is represented by the black data points. The original MC sample distribution is shown in red while the MC distributions after the application of the kinematic weights and then the kinematic + decay model weights are seen in blue and green respectively. The figure is made with the 2011+2012 dataset with the same behaviour seen for the Run 2 datasets.

100 and the learning rate, defined by the XGBoost variable η, was fixed to a value of 0.3. These values were slightly modified from the default training values as the values used were found to result in a slightly better level of performance.

A common tool to describe the performance of a classifier is the Receiver Operating Characteristic Curve, or ROC curve. This describes the level of background rejection and signal efficiency achieved for a given classifier output. The ROC curves produced when applying the trained XGBoost classifier to the 2011+2012, 2015+2016 and 2017+2018 datasets can be seen in Figure 5.10. As can be seen from these curves, a very high level of background rejection can be achieved for all three datasets, for a wide range of signal efficiencies. The final decision on which BDT requirement value to make within the analysis is discussed in Section 5.5.1.

5.4.2 XGBoost: Over-training Studies

When training a classifier it is important to ensure that the level of bias and over-training which is developed during the training process is kept to a minimum. Over-training occurs when the classifier is given too much information about the training dataset such that is recognises artificial features of the training dataset, such as statistical fluctuations, as genuine features. To a certain extent, over-training is always present within a BDT classifier due to the mechanics of the

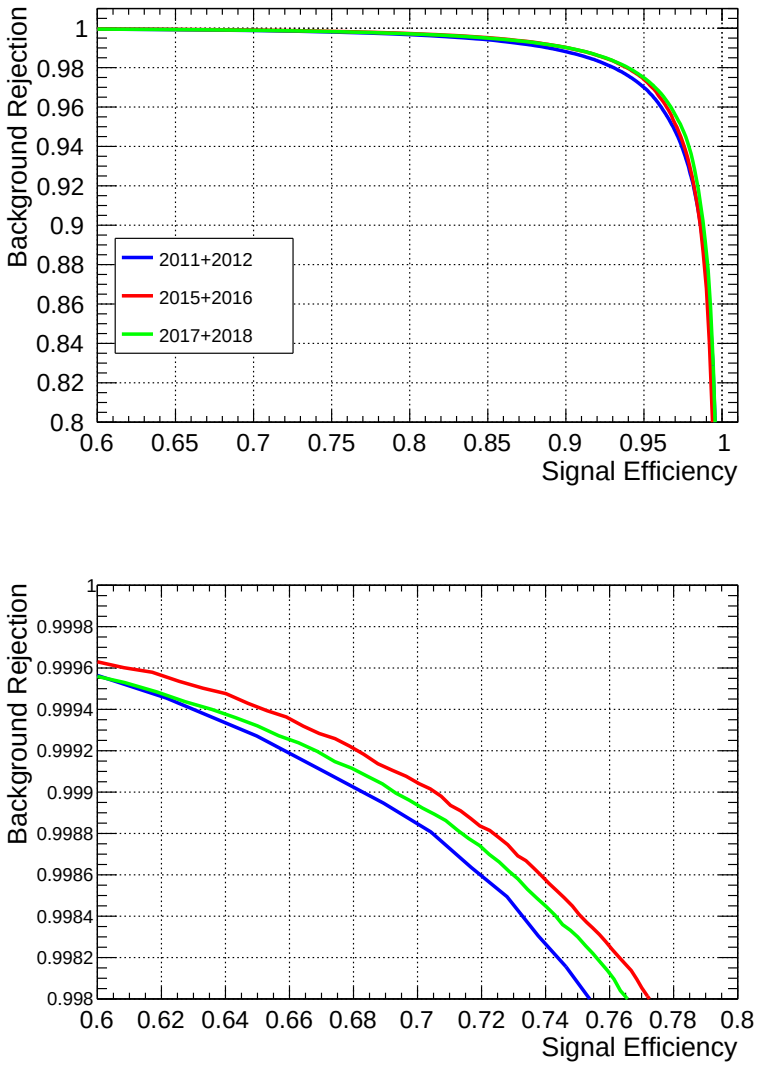

Figure 5.10: Receiver Operating Characteristic (ROC) curves for the 2011+2012 dataset (blue), the 2015+2016 dataset (red) and the 2017+2018 dataset (green). The lower plot shows the ROC curve zoomed into the signal efficiency region relevant for the analysis.

underlying classification process. However, it is imperative that this over-training is not too large and different tests can be made to ensure that the over-training levels are minimised.

Figure 5.11 shows, for the Run 1 and 2015+2016 classifier, the BDT response for the signal and background proxies, for both the training and testing samples. A large level of

over-training would present itself as a noticeable difference in the BDT distribution between the training and testing samples. As can be seen from the figure, the BDT distributions for the training and testing samples are very similar suggesting that the level of over-training present is small. The different shape of the BDT response distribution between the Run 1 and 2015+2016 datasets can be understood by considering the meaning of the BDT response variable output. The BDT output number represents the signal purity of the training sample. A BDT output closer to one implies the candidate is more likely to be a signal candidate and an output closer to zero implies the event includes a background candidate. The dependence of the BDT output on the signal purity for a given BDT output requirement means that the BDT response distribution will differ between the Run 1 and 2015+2016 training as the relative size of the two proxy samples is different between the two classifiers.

An additional tool that can be used to check the performance of the BDT, whilst ensuring that over-training is minimised, is the classifier learning curve which shows how the integrated area under the ROC curve varies as a function of a specific learning parameter, such as the number of trees used when training. The area under the ROC curve is a direct measure of the performance of the classifier and so, generally, classifiers are optimised to maximise this metric. It is possible to check the learning curves for both training and test samples simultaneously where a gradual convergence of the two learning curves indicates a lack of over-training. Figure 5.12 shows examples of learning curves for the Run 1 and 2015+2016 training. The top plots show the performance of the classifiers for the training and test samples as a function of the fraction of the total dataset used. For small data samples the level of overtraining is larger and the performance of the two samples diverges as expected. Equally, no large divergence can be seen for larger datasets which would indicate a bias. The bottom figures show the performance of the classifier for the training and testing samples as a function of the number of trees used in the training. Similarly, a large level of divergence between the two curves would indicate significant over-training. As was mentioned at the beginning of this section a small level of overtraining is unavoidable for a BDT classifier. However, the small level of divergence seen here suggests that the level of over-training present within the classifier is small enough to be neglected.

Over-training can also be checked for by performing a cross-validation test where the classifier is trained on a subset of the data before being applied on a different subset. If the classifier is unbiased then the classifier should not show significant performance improvements when using different subsets of the data for the training. Within this analysis a cross-validation test is performed by creating subsets from the different data-taking years present within the data. ROC curves for the different train-application combinations can be seen in Figure 5.13. As can be seen, the level of classifier performance variation between the different subsets is small. The variation is seen to be slightly larger for the 2015+2016 training and this is understood to be due

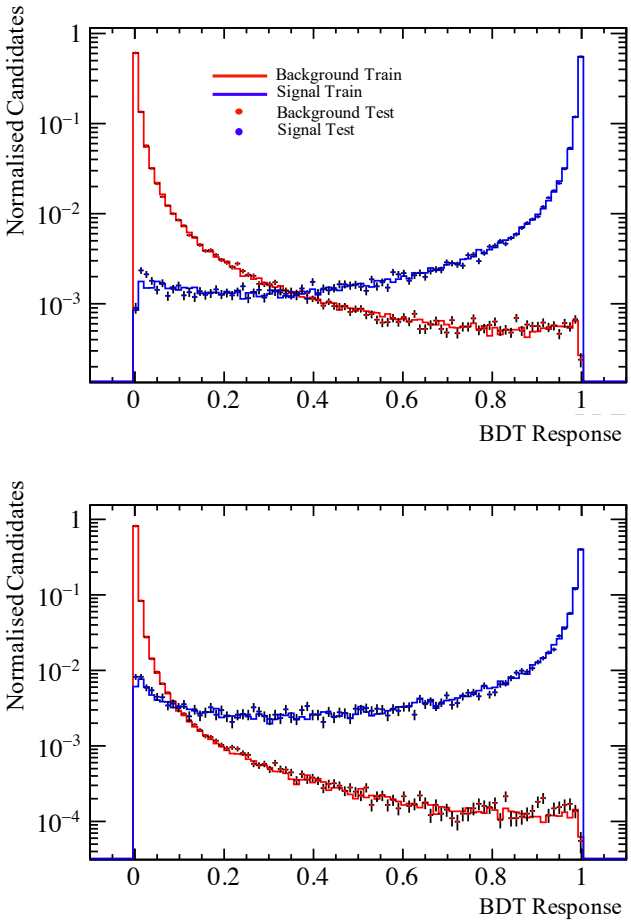

Figure 5.11: The XGBoost BDT response for the signal (blue) and background (red) samples. The training data is represented with the histograms while the test data is represented with the data points. The top plot shows the Run 1 response and the bottom plot the 2015+2016 response.

to the greater levels of variation in the number of events which make up the different 2015+2016 MC year subsets.

A final check of over-training is made to validate the BDT training. This test involves the comparison of the BDT output distributions when the BDT is applied on different MC modes which all share the same topology and, therefore, are expected to behave similarly. Figure 5.14 shows the BDT response on the $B^0 \to \pi^+\pi^-\mu^+\mu^-$, $B^0_s \to \pi^+\pi^-\mu^+\mu^-$, $B^0 \to K^+\pi^-\mu^+\mu^-$, $B^0 \to J/\psi\,\pi^+\pi^-$, $B^0_s \to J/\psi\,\pi^+\pi^-$ and $B^0 \to J/\psi\,\pi^+\pi^-$ MC samples. This is a particularly useful test on the rare mode MC samples as the decay models of these modes are not well known. Therefore, if the BDT output shows consistency between resonant and rare mode MC

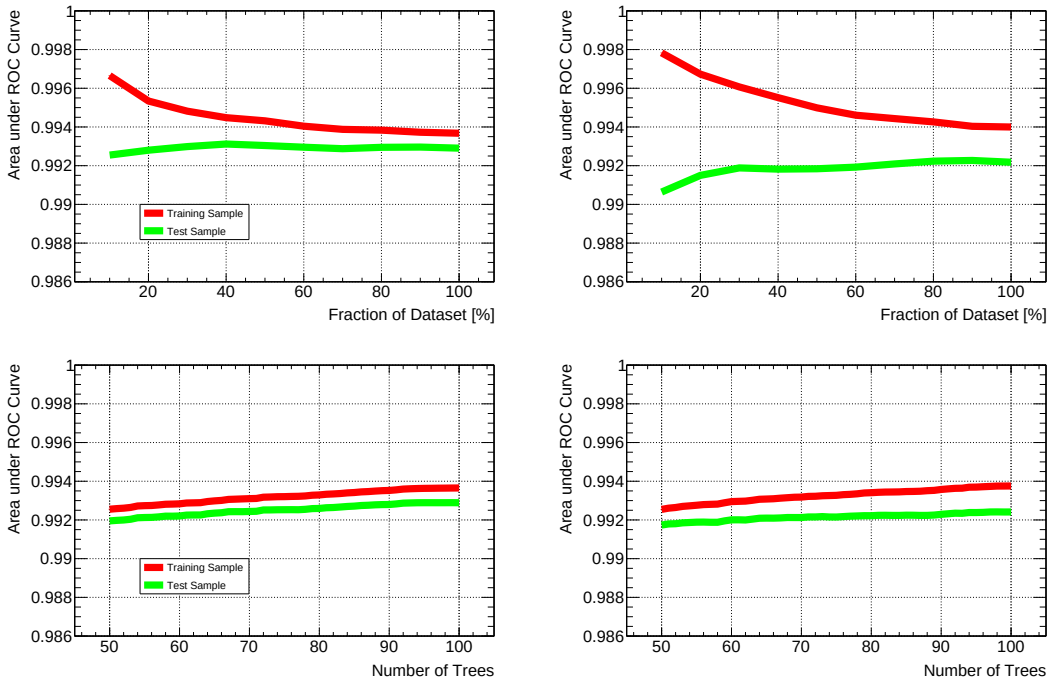

Figure 5.12: Learning curves showing the performance of the classifiers for the Run 1 training (left plots) and the 2015+2016 training (right plots). The top row shows the performance of the classifier as the fraction of the total dataset which is used to train the classifier is increased. The bottom row shows the performance of the classifier as the number of trees used within the training is increased.

samples then this adds confidence in the ability of the classifier to produce similar results despite different intermediate resonance contributions to the hadronic final-state particles. As can be seen the shapes of the BDT response between the different distributions are very similar for the 2011+2012, 2015+2016 and 2017+2018 datasets. In order to make the BDT response more meaningful the BDT response is transformed within the figure. This is necessary as the signal is normally peaked towards high values and so by transforming the variable the subtler differences between the different responses can be seen.

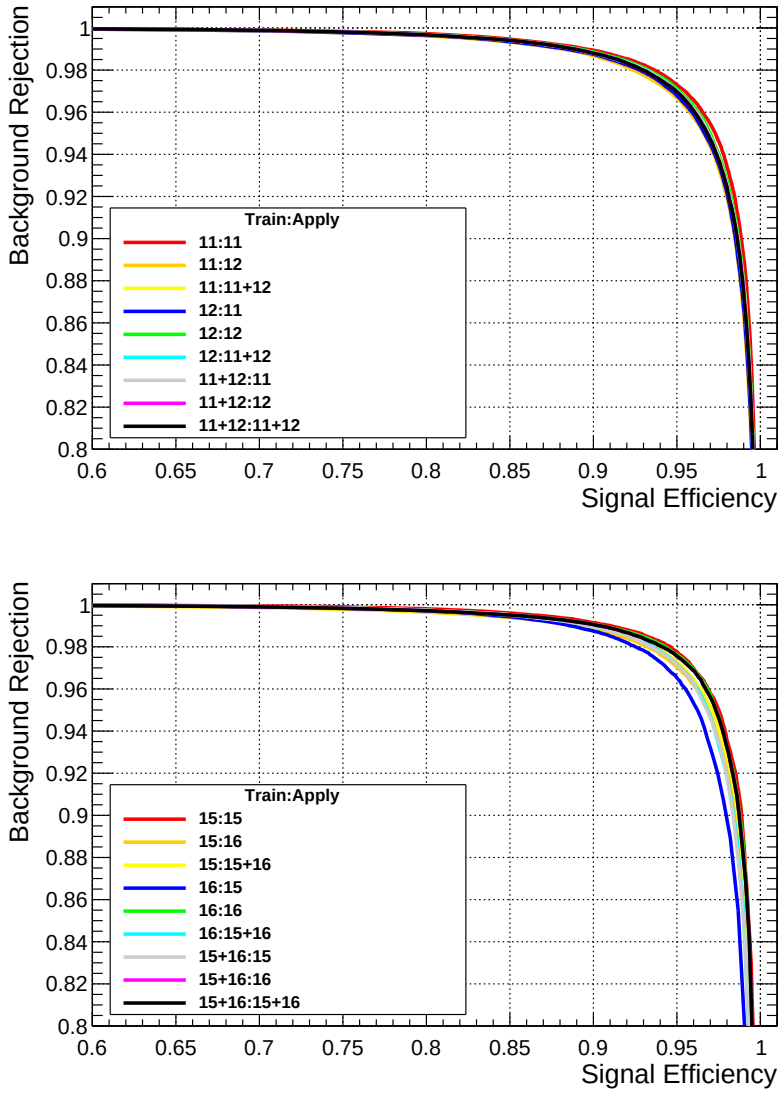

Figure 5.13: Cross-validation ROC curves which show the performance of the BDT classifier when training using a subset of the data and then applying on a different subset. The top set of ROC curves show the Run 1 training and the bottom set the 2015+2016 the training.

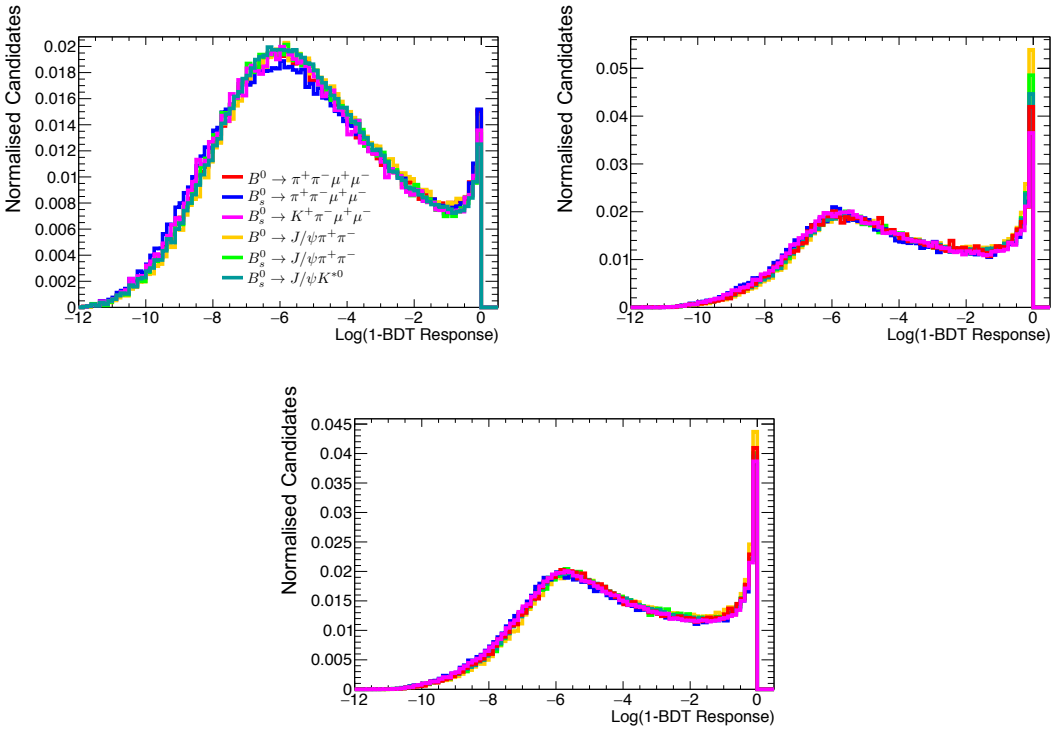

Figure 5.14: The XGBoost BDT output when applied on different MC samples which share a similar topology. The top left plot shows the distributions for the Run 1 MC samples, the top right plot shows the distributions for the 2015+2016 MC samples and the bottom plot for the 2017+2018 MC samples.

5.5 Particle Identification vs BDT Response Selection

As was previously discussed in Chapter 3 the PID information provided by the LHCb detector allows for the identification of different particle species within an event. During an off-line physics analysis the LHCb PID information can be accessed through different sets of variables. Within this analysis the PID variables used are known as the `ProbNN` variables [169]. These are the output of a neural network that is trained using data samples to give the probability that a given particle track belongs to a particular species using the LHCb RICH system, muon system, calorimetry system and ghost track probability information. The `ProbNN` variables therefore have a value between 0 and 1 and requirements on these values can be made to reduce the presence of backgrounds within an analysis which arise due to particle species mis-identification.

Within this analysis the PID requirements which are applied on the final-state hadron tracks to select out the rare and resonant modes take the form shown in Equation 5.1

$$\texttt{ProbNN}\pi \ \star(1-\texttt{ProbNN}K)\star(1-\texttt{ProbNN}p). \tag{5.1}$$

By combining the requirements in this way the probability that the final-state hadron is a pion is combined with the probabilities that it is not a kaon or proton, allowing for a more powerful rejection of K-π and p-π mis-identification backgrounds. In addition to these PID requirements, the requirements placed on the final-state muons take the form `ProbNN`μ which serves to reduce the levels of fully hadronic backgrounds present within the dataset. Finally, the response of the LHCb muon system, which is summarised by a boolean variable, `IsMuon` is applied to the final-state tracks. A particular track is given a true `IsMuon` value if it has a minimum number of hits in the muon stations. The number of required hits depends on the momentum of the track as low momentum tracks are absorbed and do not reach further downstream in the detector. Within this analysis the two pion tracks are required to return false on an `IsMuon` query, while the muons are required to return a true value.

The value of the `ProbNN` requirements made in the analysis needs to be carefully considered due to the necessity of maintaining a balance between a reduction in background levels and the retention of signal candidates. Within the analysis, the final choice of working point for both the PID requirements and the BDT response selection is made simultaneously by optimising the expected statistical significance of a $B^0 \to \pi^+\pi^-\mu^+\mu^-$ observation with a pseudo-experiment study. The pseudo-experiment study (which is discussed in detail in Section 5.5.1) makes use of PID efficiency estimates. However, as was the case in Section 5.3.1 where the standard LHCb MC samples fail to accurately model certain kinematic variables, the PID variables present within the MC samples are known to also disagree with data. This is due to the sensitivity

Table 5.4: PID Fiducial requirements for Run 1 and Run 2

Dataset	Variable	Requirement
Run 1 + Run 2	All Tracks	$3 < p < 100\,\text{GeV}/c$ $1.5 < \eta < 5$ $p_\text{T} > 250\,\text{MeV}/c$
Run 1	Occupancy	$0 <$ nTracks < 500
Run 2	Hadrons	$p \notin [5.5, 11.5]\,\text{GeV}/c$ and $\eta \notin [4.1, 5.0]$ $p \notin [3.0, 5.5]\,\text{GeV}/c$ and $\eta \notin [3.75, 4.1]$ $p \notin [3.0, 5.5]\,\text{GeV}/c$ and $\eta \notin [3.45, 3.75]$
	Occupancy	$0 <$ nSPDHits < 600

of the PID variable responses to factors within a specific event, such as the detector occupancy. In order to calculate the correct PID efficiencies, the LHCb `PIDCalib` package is used [170]. This package uses calibration samples taken from data that contain decays where a particle species hypothesis can be determined using kinematic information only, as was discussed in Chapter 3. By binning both the calibration samples and the MC samples in kinematic and occupancy variables, the corresponding PID efficiency from the calibration sample can be looked up and used as the PID efficiency for each MC candidate.

Within this analysis the `PIDCalib` calibration samples are binned in the track momentum, track pseudo rapidity and detector occupancy, the latter being represented by the number of tracks in Run 1 and the number of SPD hits in Run 2. However, prior to using the `PIDCalib` package it is necessary to make sure that the phase-space coverage of the calibration samples matches that of the MC samples. This is necessary because otherwise candidates within the MC samples would receive a meaningless PID efficiency. These candidates are first removed removed via the application of a series of fiducial requirements that are listed in Table 5.4. For consistency, these fiducial requirements are also applied to the Run 1 and Run 2 datasets.

5.5.1 Working Point Optimisation

To perform the optimisation study 1000 pseudo experiments are generated and fit with a model in order to evaluate the average significance of a $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal for different combinations of PID and BDT requirements. This procedure, which is performed separately for the three datasets and across the full q^2 range, calculates the significance of each pseudo experiment, at each working point, using Wilk's theorem [171]. Using Wilk's theorem the statistical significance, S, can be expressed as

$$S = -2\big(\ln(L_1) - \ln(L_0)\big) \tag{5.2}$$

where L_1 is the likelihood calculated under the signal hypothesis and L_0 is the likelihood calculated under the null hypothesis. To make the different BDT working points differentiable from one another, the BDT response at this stage of the analysis is flattened whereby the raw BDT output is replaced with a corresponding BDT efficiency.

The model used within the optimisation procedure is fit, using a maximum likelihood fit, to the invariant $m(\pi^+\pi^-\mu^+\mu^-)$ mass using a set of probability density functions (PDFs). The same PDFs are used in both the generation and fitting of the pseudo-experiment data with the form of the PDFs used discussed in Chapter 8. The PDF shapes for the components are determined from separate fits to MC samples with three PDF components included corresponding to the decays $B^0 \to \pi^+\pi^-\mu^+\mu^-$, $B^0_s \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ where the kaon is mis-identified as a pion by setting the invariant mass of the kaon track to the known pion mass. Finally a final PDF is included to describe the combinatorial background component. When applying the PID variable requirements on the data, the shapes of the signal and background components within the data is altered. This shaping is larger for a background component that results from the PID mis-identification of particle tracks, such as $B^0 \to K^+\pi^-\mu^+\mu^-$. In order to include this effect within the optimisation the PID efficiencies from the `PIDCalib` package are applied as weights to the different MC samples before they are fit to determine the different PDF parameters. The extent of the shape alteration differs between working points meaning that these fits are repeated for each set of PID and BDT requirements.

The combinatorial exponential lineshape is modelled using an exponential distribution from a separate fit to the upper sideband of the data. The exponential parameter is allowed to float within each pseudo experiment with the initial value set from the initial upper sideband fit.

The number of events generated within the toy datasets is determined by first calculating the expected yields of the different components. The yield for the combinatorial component is determined by extrapolating the yield from the upper sideband fit to the full fit range. This is repeated for each separate working point. The yields for the $B^0(B^0_s) \to \pi^+\pi^-\mu^+\mu^-$ components are determined by scaling the yield from the control mode fit of Chapter 8 using ratios of efficiencies and branching fractions, following Equation 4.2 from Chapter 4. The yield for the $B^0 \to K^+\pi^-\mu^+\mu^-$ background is calculated from a separate fit to the data where alternative PID requirements are applied alongside a mass window selection around the $K^{*0}(892)$ resonance peak to determine the yield of $B^0 \to K^{*0}\mu^+\mu^-$ events which dominate the $B^0 \to K^+\pi^-\mu^+\mu^-$ background. This yield is then scaled with `PIDCalib` and mass requirement efficiencies. The yield of the $B^0 \to K^+\pi^-\mu^+\mu^-$ component in the pseudo experiments is then constrained in the fit with a Gaussian function which takes the scaled yield and uncertainty as mean and width parameters.

An example of a single fitted toy dataset can be seen in Figure 5.15. After the calculation

of the significance of each pseudo experiment using Wilk's theorem, the final significance for the corresponding PID and BDT working point is determined from the mean of the resulting Gaussian distribution of the 1000 individual toys experiments.

5.5.2 Optimisation Results

The results of the optimisation procedure can be seen in Figure 5.16. The plots shown correspond to studies performed with muon PID requirements of $\texttt{ProbNNmu} > 0.10$. The process was repeated with an alternative muon PID requirement of $\texttt{ProbNNmu} > 0.05$, however, this was found to result in worse performance. For all three of the datasets, a clear area of larger statistical significance can be seen. Within these regions many of the adjacent working points yield a similar expected significance. Therefore, the final working points for the three datasets were ultimately chosen by also considering the signal and background yields, for each working point, which were calculated for the study. As a tighter set of requirements results in a smaller signal yield, the final working points were not selected to be the tightest requirements available within the significant region. Instead, the chosen requirements reach a compromise between lying within the optimal region and being loose enough to ensure a larger signal yield.

The optimal requirements are selected to be

$$\texttt{ProbNN}\pi \ \star\texttt{(1-ProbNN}K\texttt{)}\star\texttt{(1-ProbNN}p\texttt{)} > 0.6 \tag{5.3}$$

for all three datasets with $\texttt{ProbNNmu} > 0.10$ requirements on the muons and a flattened BDT requirement corresponding to a 60% BDT efficiency.

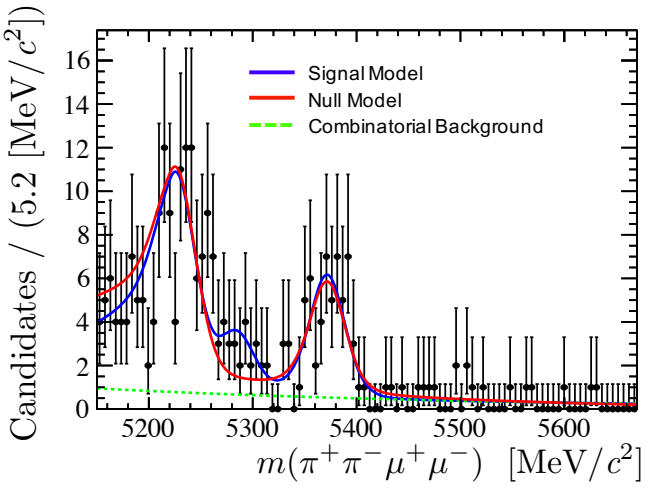

Figure 5.15: Example of a single pseudo-experiment where the generated dataset is fit with a model to determine the statistical significance of a $B^0 \to \pi^+\pi^-\mu^+\mu^-$ observation. In this example only a loose requirement is made on the PID variables and a large $B^0 \to K^+\pi^-\mu^+\mu^-$ is present.

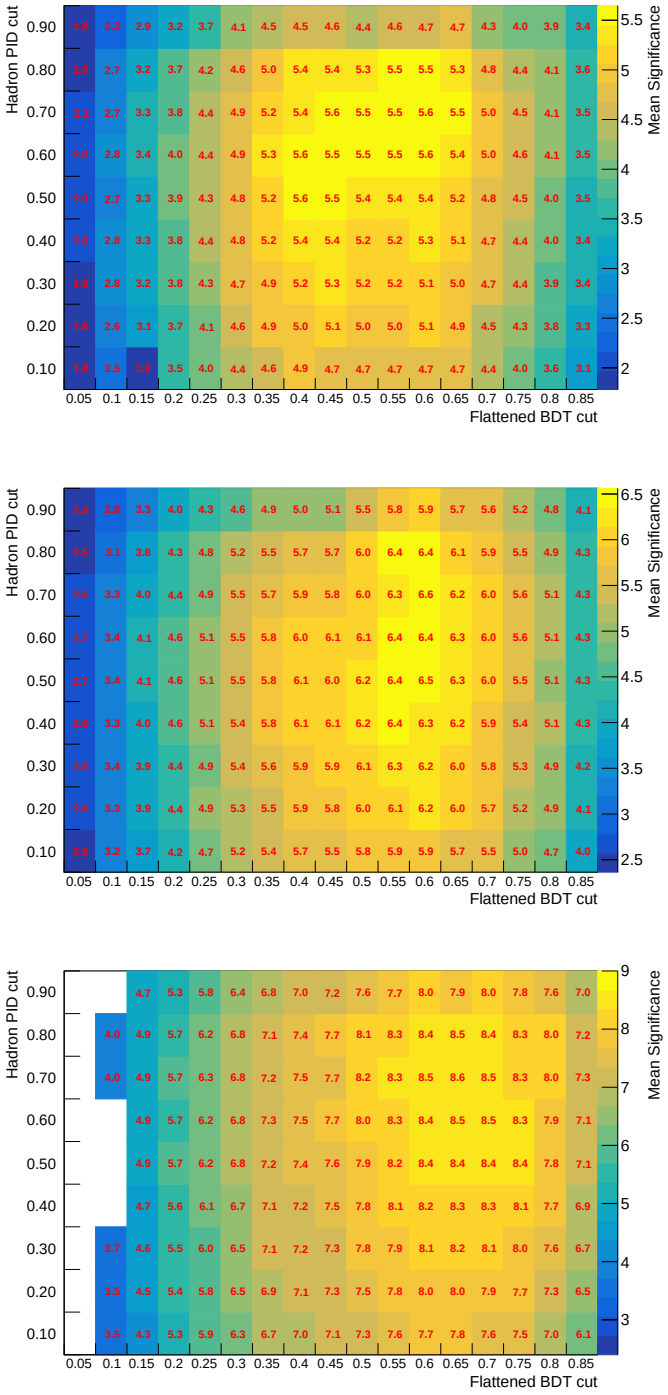

Figure 5.16: PID vs BDT working point optimisation results for Run 1 (top), 2015+2016 (middle) and 2017+2018 (bottom). The results show the mean significance of a $B^0 \to \pi^+\pi^-\mu^+\mu^-$ observation for each PID and BDT working point.

5.6 q^2 vs $m(\pi^+\pi^-)$ Selection

At this stage of the analysis, with the majority of the selection procedure completed, the possibility of performing the analysis in different bins of $m(\pi^+\pi^-)$ and q^2 is explored. As was discussed within Chapter 2, being able to provide measurements of the rare mode branching fractions in different regions of q^2 opens up sensitivity to different classes of new physics and is more comparable to different theoretical predictions. Furthermore, the possibility of performing the analysis in bins of $m(\pi^+\pi^-)$ is also considered and is of interest given the wide widths of the ρ^0 and $f_0(980)$ resonances and the presence of the higher mass states which are seen in the $m(\pi^+\pi^-)$ spectra within previous LHCb analyses [172, 173].

In order to determine the practicality of making binned branching fraction measurements, a $m(\pi^+\pi^-)$ versus q^2 binned study is performed whereby the yields for the signal, the combinatorial background and the dominant K-π particle mis-identification background are estimated in the different bins. The main purpose of this study is to determine whether, after the application of the analysis selection, the predicted signal yield for the rare B^0 mode is high enough to be measured in each bin separately.

For the study, three q^2 bins corresponding to [0-8] GeV^2/c^4, [11-12.5] GeV^2/c^4 and [15-19] GeV^2/c^4 alongside $m(\pi^+\pi^-)$ bins corresponding to [0.5-1.3] GeV^2/c^4 and [1.3-2.0] GeV^2/c^4 are used. The boundary of the lower $m(\pi^+\pi^-)$ bin is chosen to mimic that of a previous LHCb analysis [122]. The optimal BDT, fiducial and PID requirements are made and, in addition, $m(\pi^+\pi^-\mu^+\mu^-)$ mass requirements corresponding to the signal region of $m(\pi^+\pi^-\mu^+\mu^-) \in [5170, 5800]\,\mathrm{MeV}/c^2$.

The combinatorial background yield is determined by counting the number of events in the upper sideband of the data in each bin. The low number of events in the upper sideband after the application of the selection means that fitting an exponential curve to the upper sideband to extrapolate the yield to the signal region is not possible so this yield is left unscaled.

Both the $B^0 \to K^+\pi^-\mu^+\mu^-$ background and $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal yields are estimated in the lower $m(\pi^+\pi^-)$ bin by scaling the control mode yield from Chapter 8 using the ratio of branching fractions and efficiencies as is done in Equation 4.2. However, this estimation method is only used for the lower $m(\pi^+\pi^-)$ bin because the branching fractions assumed in the calculation were calculated, in separate, previous analyses, within this range. To estimate the yield in the upper $m(\pi^+\pi^-)$ bin, the yield in the lower bin is multiplied by an additional factor that accounts for the fraction of events expected in the upper bin, relative to the lower bin. This factor is estimated using the decay model described in Chapter 6.

The resulting estimated yields can be seen in Figure 5.17 where the given yields are summed over all three datasets. Due to the low expected yield in the higher $m(\pi^+\pi^-)$ bin, the

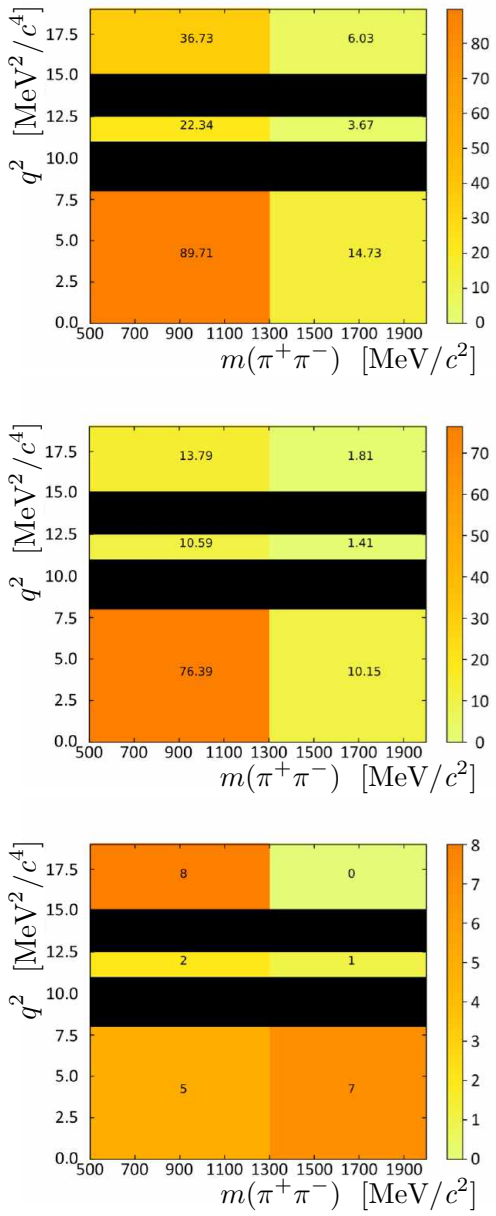

Figure 5.17: Estimated yields for the combined Run 1 and Run 2 dataset in bins of q^2 and $m(\pi^+\pi^-)$ for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ mode (top), the background decay $B^0 \to K\pi\mu^+\mu^-$ (middle) and the number of events within the upper sideband of the data attributed to combinatorial background (bottom). The black bands show the vetoes applied to remove the charmonium resonances.

analysis proceeds using the lower $m(\pi^+\pi^-)$ bin. In addition, the lower $m(\pi^+\pi^-)$ bin boundary is raised within the analysis from $0.5\,\mathrm{GeV}/c^2$ to $0.515\,\mathrm{GeV}/c^2$ to effectively veto any $\pi^+\pi^-$ backgrounds originating from the decay of a K^0_{S} meson.

From the study it is also concluded that the branching fraction can only be measured in the low and high q^2 bins, and integrated across the three q^2 bins. The possibility of performing a measurement in the middle q^2 bin is ruled out due to the smaller number of predicted events in that bin. This region is also of less interest due to the difficulty in making accurate theoretical predictions in this q^2 range.

5.7 Peaking Backgrounds

In addition to combinatorial backgrounds, other sources of background are present within the data which are reconstructed within the invariant mass fits as peaking structures. These peaking backgrounds can lie in the same mass region as the rare mode signal peaks and this close proximity can make separation of the signal and peaking backgrounds difficult. Furthermore, many of these peaking backgrounds share a similar topology with that of the signal mode which means that the BDT classifier does not remove peaking backgrounds as efficiently as it removes combinatorial background. The identification of peaking backgrounds is therefore a key consideration within the analysis selection procedure and those backgrounds identified as dangerous to the analysis can be reduced, or even removed, using combinations of PID requirements and invariant mass vetoes.

In this analysis the different classes of peaking backgrounds considered include over reconstructed backgrounds, partially reconstructed backgrounds and backgrounds arising from the the mis-identification of particle species. More specific examples of these primary types of peaking background are discussed within the following sections with some examples being fully hadronic backgrounds, charm backgrounds and pion-muon swap backgrounds. Many of these peaking backgrounds are relevant for both the rare modes and the resonant modes. This means that the resonant modes, where the peaking backgrounds are often far easier to identify, can be used as a proxy to identify the backgrounds that will also be relevant for the rare mode. This technique for identifying relevant peaking backgrounds has been used previously in other rare decay searches, such as when determining the branching fractions of the $B^0_s \to \overline{K}^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ processes [91, 174]. Noticeable exceptions where this technique cannot be used include the identification of semi-leptonic cascade decays, swap backgrounds and fully hadronic decays which are only relevant for the rare modes because their contribution to the resonant modes is so small.

5.7.1 Identifying Backgrounds

There are several different techniques that can be used to identify the presence of peaking backgrounds within the data. The most straightforward of these is to identify the signature visually by applying a set of requirements to the dataset which are designed to accentuate the specific background above other features of the invariant mass distribution. In this analysis, for many peaking backgrounds that originate from particle species mis-identification, this is done by placing a set of loose `ProbNN` requirements on the final-state particles which intentionally select out background. This is then complemented by reconstructing the invariant mass combinations of the final-state tracks under alternative mass hypotheses. Other peaking backgrounds, which do not originate from a mis-identification, can also be seen by reconstructing the invariant mass of a subset of the final-state particles and by applying loose PID requirements in addition to the optimal BDT requirement.

Not all of the peaking backgrounds of relevance to the analysis can be identified in this way. The relevance of other peaking backgrounds to the analysis can be determined by estimating the yield of these backgrounds within the final invariant mass fits. The expected yields of these backgrounds are estimated using a combination of branching fraction information, efficiencies and simulated samples as

$$N_X = N_{J/\psi K^{*0}} \times \frac{\epsilon_X}{\epsilon_{J/\psi K^{*0}}} \times \frac{\mathcal{B}(X)}{\mathcal{B}(B^0 \to J/\psi K^{*0}) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-) \cdot \mathcal{B}(K^{*0} \to K^+\pi^-)} \times \frac{f_q}{f_d}\,, \tag{5.4}$$

where N_X represents the yield of the background mode, X; $N_{J/\psi K^{*0}}$ represents the control mode yield; ϵ represents the product of the q^2, fiducial requirement, mass requirement and PID efficiencies; $\mathcal{B}$ is the relevant branching fraction and $\frac{f_q}{f_d}$ represents a fragmentation fraction ratio. For this exercise all branching fractions are taken from the world-values provided from the Particle Data Group (PDG) tables [175] and, where these branching fractions are absent, the values are estimated using ratios of similar decays.

After the calculation of the estimated yields, a conclusion can be drawn on whether the predicted yield is large enough to be considered problematic. Unfortunately, the number of peaking backgrounds which are considered for the analysis is too large to ensure that each of these backgrounds can be modelled using the full LHCb simulation. Therefore, the simulated samples used for the yield calculation exercise are generated using the `RapidSim` software package.

Many backgrounds are effectively removed following the application of the full analysis selection and so these backgrounds need not be considered further. Backgrounds which remain

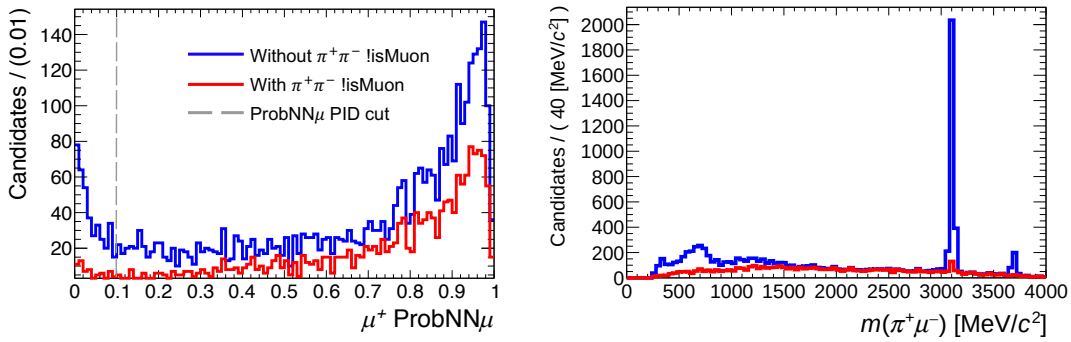

Figure 5.18: Left: The `ProbNN`μ distribution for the positive muon track taken from the Run 1 dataset, before and after the application of the `!IsMuon` on the final-state pions. Right: The reconstructed $\pi^+\mu^-$ invariant mass for the Run 1 rare mode where large pion-muon swap backgrounds can be seen before and after the `!IsMuon` pion condition.

post-selection need to be included within the final mass fits, or, if they are considered small enough, can be treated as a source of systematic uncertainty within the analysis.

Yield estimates for different classes of peaking background can be found in Table 5.5 at the end of the chapter where, for the vast majority of identified backgrounds, the estimated yields are small enough for the peaking background to be safely disregarded. The rest of this chapter discusses the different classes of background considered and, where applicable, demonstrates the existence of the peaking backgrounds from the datasets. The figures included, unless stated otherwise, are created using the Run 1 dataset. The conclusions drawn are the same for all three datasets however as the same background features can be reproduced from any of the datasets.

5.7.2 Hadronic Decays

Fully hadronic peaking backgrounds are highly suppressed within the analysis by application of the `IsMuon` condition, at stripping level, on the final-state muon tracks. However, given that there is an approximately 2% mis-identification efficiency per track for the `IsMuon` condition [162], potential backgrounds which have a branching fraction of order 10^{-5} or greater may prove to be potentially dangerous given that the expected rare mode branching fractions are of the order 10^{-8}.

The presence of hadronic backgrounds can be checked for with the data by examining the `ProbNN`μ distribution of either of the final-state muon tracks. As can be seen in the left plot of Figure 5.18, a peak can be seen in this distribution close to zero which indicates a low level of confidence in the hypothesis that the track is in fact a muon. This peak is reduced by

requiring that the final-state pion tracks return a negative `IsMuon` condition from the LHCb muon system -- demonstrating that some of these cases are caused by the simultaneous incorrect identification of a pion and a muon, discussed further in the next section. To further improve the rejection of these backgrounds a PID `ProbNN`μ requirement of > 0.1 is made to the dataset, as was mentioned previously.

5.7.3 π-μ Swap Backgrounds

The right plot of Figure 5.18 shows the presence of backgrounds which arise when a muon is mis-identified as one of the final-state pions and vice versa. The background can be seen visually in the $m(\mu^+\mu^-)$ mass distributions after swapping the μ^+ and π^+ or the π^- and μ^-. Clear contributions are seen from both J/ψ and $\psi(2S)$ meson decays. This background is seen to be heavily suppressed with the application of the `!IsMuon` requirements on the final-state pions and the small level of background which survives the implementation of the requirement is then vetoed by rejecting events where either $m(\pi^+\mu^-)$ or $m(\pi^-\mu^+) \in$ [3076,3116] MeV/c^2.

5.7.4 Mis-identified Backgrounds

Peaking backgrounds which originate from the incorrect identification of final-state particle species comprise some of the largest backgrounds for the analysis. These backgrounds are heavily reduced in the data with the application of the selection PID requirements, however, some backgrounds are seen to survive.

Backgrounds originating from $K \to \pi$ mis-identification, where a kaon from the final-state is mis-identified as a pion provides the largest source of these types of background with $B^0 \to K^{*0}(892)\mu^+\mu^-$ and $B^0 \to K^*(1430)^0\mu^+\mu^-$ the most prominent examples. An example of the resonant version of these backgrounds can be seen in the left plot of Figure 5.19, which shows the reconstructed dihadron mass under a $(\pi \to K)\pi$ mass hypothesis and with loose reversed PID requirements applied. As can be seen two large peaks, corresponding to the K^{*0}(892) and the $K^*(1430)^0$ are visible and although these peaks are heavily suppressed by the application of the full selection, some broad background structures remain. These backgrounds are particularly problematic for the analysis as they lie almost on top of the predicted signal peak. This means that the background cannot be safely vetoed and therefore the background needs to be included within the final resonant and rare fits.

In addition to a single $K \to \pi$ identification it is also possible to have a double $K \to \pi$ identification. A double mis-identification of this kind would lead to the existence of background from decays such as $B_s^0 \to \phi\mu^+\mu^-$ within the dataset. By reconstructing the dipion final-state under a $(\pi \to K)(\pi \to K)$ mass hypothesis and by again applying loose reversed PID

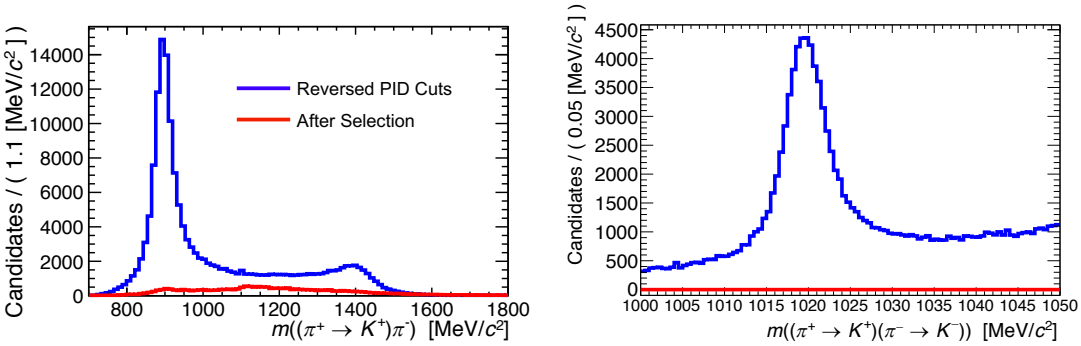

Figure 5.19: Left: The $m((\pi^+ \to K^+)\pi^-)$ mass distribution taken from the Run 1 dataset showing the presence of kaon-pion PID backgrounds before and after the signal selection. Right: The double kaon-pion PID background originating from a $\phi(1020)$ resonance seen in the Run 1 dataset which is seen to be removed after the application of the full selection. Both plots are made using candidates which contribute to the resonant mode datasets.

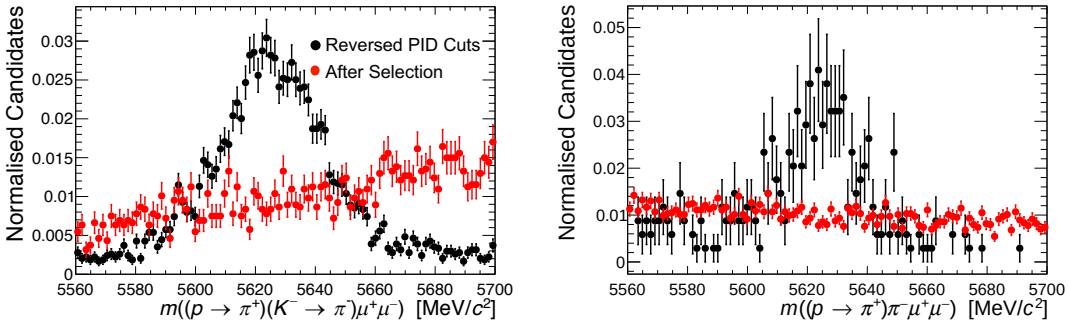

Figure 5.20: Peaking backgrounds from $\Lambda_b^0 \to pK^-\mu^+\mu^-$ (left) and $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ (right) decays seen in the Run 1 dataset. The backgrounds can be seen to be removed after the application of the full selection.

requirements, a clear peak in the dihadron spectrum becomes visible, as seen in the right plot of Figure 5.19 for the resonant mode. The tight requirements placed on the pion `ProbNN` variables as part of the selection lead to this background mode being double-suppressed and the background is seen to be completely removed from the datasets after the PID requirements are imposed.

Another source of mis-identified background stems from the identification of a final-state proton as a final-state pion, with $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays that contain an additional kaon-pion identification a prominent example. This background is identified by reconstructing $m((\pi^+ \to p)(\pi^- \to K^-)\mu^+\mu^-)$ within the data and applying loose reversed PID requirements

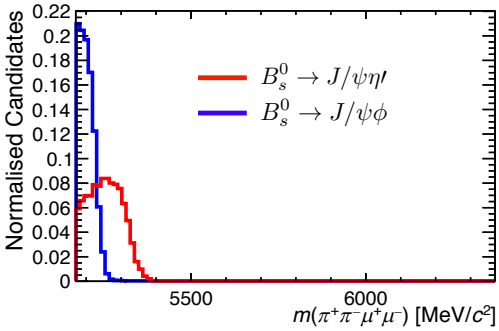

Figure 5.21: Run 1 MC samples showing the shapes of the $B_s^0 \to J/\psi\,\eta'$ (Red) and $B_s^0 \to J/\psi\,\phi$ (Blue) backgrounds in the resonant mode signal window.

as is done for the other mis-identification backgrounds. A clear background peak originating from a Λ_b^0 is seen in the left plot of Figure 5.20. Once again this background peak is seen to be completely absent from the dataset following the application of the full selection.

Decays from $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ also form a potential peaking background from when a proton is mis-identified as a pion. This background is identified in a similar way to the other Λ_b^0 background. It can be seen that this background appears smaller than the background from the $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decay. This is due to the branching fraction of the $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ mode being an order of magnitude smaller than that for $\Lambda_b^0 \to pK^-\mu^+\mu^-$ [176, 177]. Once again, this background is seen to be entirely removed from the dataset after the application of the full selection, as can be seen in the right plot of Figure 5.20.

5.7.5 Partially Reconstructed Decays

Decay modes which lead to a set of final-state particles that contain $\pi^+\pi^-\mu^+\mu^-$ as a subset can form broad structures in $m(\pi^+\pi^-\mu^+\mu^-)$ in the analysis where one of the additional particles is not reconstructed. These peaking backgrounds are shifted to lower invariant masses and also smeared out within the invariant mass distributions. In addition to decay modes where a charged pion or kaon is not reconstructed, many potential decays exist where a neutral pion or photon is missing and these decays are hard to identify within the invariant mass distributions due to large variations in the four-momenta of the missing particle. For the rare mode, there is the additional consideration of semi-leptonic cascade decays from an open charm meson which have neutrinos in the final-state that cannot be detected by the LHCb detector.

By calculating predicted yields, the most dangerous partially reconstructed backgrounds for the resonant mode are found to be from B_s^0 decays to an η' or ϕ meson. As a result,

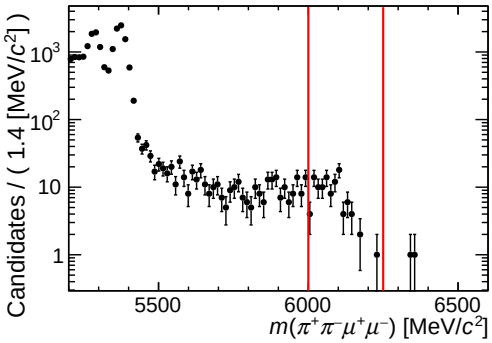

Figure 5.22: Invariant mass distribution from the Run 1 resonant mode dataset with the full PID, BDT and requirement selection applied. The roll-over seen within the highlighted region is attributed to the presence of partially reconstructed B_c^+ decays.

$B_s^0(B^0) \to J/\psi\eta'$ and $B_s^0 \to J/\psi\phi$, where $\eta' \to \rho^0\gamma$ and $\phi \to \pi^+\pi^-\pi^0$, are included as peaking backgrounds within the final resonant fit. The predicted shape of these backgrounds, taken from MC samples, can be seen in Figure 5.21. The equivalent decay $B^0 \to J/\psi\phi$ which instead originates from a B^0 mother is expected to lie outside of the signal region due to the mass shift. This background can therefore be neglected.

For the rare mode the most dangerous backgrounds arise again from a B_s^0 decaying into a η' meson with a photon which is not reconstructed. The branching fraction for this mode is currently unknown and so the $B_s^0 \to \eta'(\to \rho^0\gamma)\mu^+\mu^-$ branching fraction is estimated under the assumption that

$$\mathcal{B}(B_s^0 \to \eta'\mu^+\mu^-) \approx \mathcal{B}(B_s^0 \to \eta' J/\psi) \times \frac{\mathcal{B}(B_s^0 \to \phi\mu^+\mu^-)}{\mathcal{B}(B_s^0 \to \phi J/\psi)}, \tag{5.5}$$

which is considered to be a good approximation due to the similar topologies (see Figure 5.1). The number of estimated events for the $B_s^0 \to \eta'(\to \rho^0\gamma)\mu^+\mu^-$ background, although non-zero, is still considered to be small enough to exclude from the final invariant mass fits. Instead, the effect of including this background in the final rare mode mass fits is evaluated as a potential source of systematic error on the final result.

Backgrounds from B_c^+ decays to a $J/\psi\pi^+\pi^-\pi^+$ ($\pi^+\pi^-$) final-state are a source of partially reconstructed background that are identifiable for the resonant mode invariant mass fits. Due to the background only being present within the m($J/\psi\pi^+\pi^-$) spectrum with the loss of one (or three) pions, the background appears very smeared out and is relatively flat within the signal region. The presence of the background is verified by looking at the m($J/\psi\pi^+\pi^-$)

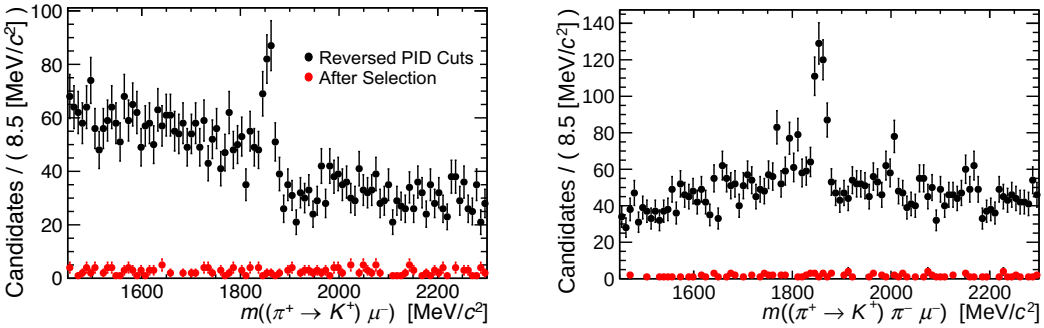

Figure 5.23: Reconstructed $m((\pi^+ \to K^+)\pi^-\mu^-)$ (left) and $m((\pi^+ \to K^+)\mu^-)$ (right) invariant masses with a small peak around the $D^0/D^\pm$ mass region seen. The background is seen to be removed after the application of the full selection requirements.

distribution at high invariant masses where a roll-over at around 6000 MeV/c^2 (as seen within the highlighted region) is attributed to the phase space of the B_c^+ meson decays. The existence of the B_c^+ backgrounds justifies the choice of $m(\pi^+\pi^-\mu^+\mu^-)$ invariant mass window used for the resonant mode fits of Chapter 8. Without the presence of this rollover within the mass fit, the B_c^+ component, which is very flat for the majority of the signal region, competes with the combinatorial background component leading to potential fit instability.

5.7.6 Charm Decays

Cascade $b \to c \to s$ decays are a specific form of partially reconstructed peaking backgrounds that are expected to impact the rare mode fits due to the relatively large rate of decays which undergo these transitions. At low invariant masses semi-leptonic $b \to c$ processes are expected to contribute through cascade decays such as $B^0 \to D^-\mu^+\nu$ where $D^- \to \overline{K}^{*0}\mu^-\overline{\nu}$ and where a K^+ is mis-identified as a π^+.

By applying reversed PID requirements, the existence of these backgrounds can be verified by reconstructing the $m((\pi^+ \to K^+)\pi^-\mu^-)$ and $m((\pi^+ \to K^+)\mu^-)$ invariant masses. This is done in Figure 5.23 using the combined Run 1 and 2015+2016 dataset as, unlike for all other peaking backgrounds considered, the shape cannot be clearly seen from just the Run 1 dataset due to the decays being largely rejected by the BDT requirement alone. Small and narrow peaks corresponding to open charm resonances are visible however after the application of the full signal selection these backgrounds are seen to be completely removed. This background is also completely removed following the application of the selection to the 2017+2018 dataset.

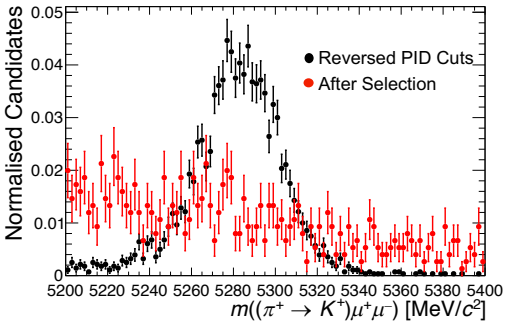

Figure 5.24: The $m((\pi^+ \to K^+)\mu^+\mu^-)$ invariant mass spectrum from the Run 1 dataset for the resonant mode where the pion track is assigned the mass of a kaon. The $B^+ \to J/\psi K^+$ peaking background is clearly visible. After the application of the selection requirements the background is removed.

5.7.7 Over Reconstructed Decays

Over reconstructed decays occur when a random particle from the event is reconstructed alongside the final-state particles which originate from a b-hadron hadron decay. This results in a background that is smeared towards higher invariant mass values. Within this analysis this class of decays could appear in the upper sideband of the signal region, where they could potentially affect the fit to the combinatorial background component.

Estimated yields for some of the potentially dangerous over reconstructed backgrounds are given in Table 5.5. These yields are calculated using the control mode fits of Chapter 8. From these fits the yield of $B^+ \to J/\psi K^+$ decays is taken and then is scaled using the mass and PID efficiencies from Chapter 7 to calculate the expected yield in the resonant mode fits. Then, the $B^+ \to J/\psi \pi^+$ background and the equivalent rare mode background yields are estimated by scaling further with branching fraction information.

The decay $B^+ \to J/\psi K^+$, where a K^+ is mis-identified as a π^+ and where a π^- is reconstructed from elsewhere within the event can be seen within the data directly by reconstructing the invariant $m(J/\psi(\pi^+ \to K^+))$ mass and applying a reverse a set of reversed PID requirements. After the application of the selection PID and BDT requirements this background is removed as can be seen within Figure 5.24. This is also seen for the rare mode sample and this background can be safely disregarded for both the resonant and rare mode fits. However, when performing the fit to the invariant mass distribution of the $B^0 \to J/\psi K^{*0}$ control mode, a set of PID requirements are made to select a $K^+\pi^-\mu^+\mu^-$ final-state combination. Therefore, although this over reconstructed background is not considered a problem for the resonant and rare mode

fits it is dangerous for the control mode and it is therefore included within the final control mode fit.

5.7.8 Summary of Peaking Backgrounds

To conclude, different sources of peaking backgrounds are considered for the resonant mode, the rare mode and the control mode invariant mass fits. These backgrounds are either identified visually within the data, or have predicted yields shown in Table 5.5 which are too large to be ignored. For the resonant mode, potentially dangerous backgrounds include $B_s^0(B^0) \to J/\psi\,\eta'$, $B_s^0 \to J/\psi\,\phi$, $B_c^+ \to J/\psi\,3\pi(\pi^+\pi^-)$ and the dominant background, $B^0 \to J/\psi\,K^+\pi^-$. All of these backgrounds are therefore included within the resonant mode mass fits. For the $B^0 \to J/\psi\,K^{*0}$ control mode fit, the over-reconstructed decay $B^0 \to J/\psi\,K^+$ is determined to be an important background and is included in the fits.

Finally, for the rare mode, the $B^0 \to K^+\pi^-\mu^+\mu^-$ mode is a major source of background however most other backgrounds which are relevant for the resonant mode are not considered dangerous. The potential exception is the $B_s^0 \to \eta'\mu^+\mu^-$ decay, whose inclusion in the final fit is evaluated as a systematic uncertainty.

5.8 Final Selection Requirements

The final set of selection requirements that are used to select the candidates for the control mode, resonant mode and rare mode invariant mass fits can be found in Table 5.6. The requirements within the table are applied to the data in addition to the stripping requirements from Table 5.2, the trigger requirements from Table 5.1 and the fiducial requirements of Table 5.4.

Background Class	Decay Mode	B	2011+2012	2015+2016	2017+2018
Backgrounds	$B^0 \to \rho^0\rho^0$	9.6×10^{-7}	~ 0	~ 0	~ 0
	$B^0 \to K^{*0}(\to K^+\pi^-)\pi^+\pi^-$	5.5×10^{-5}	~ 0	~ 0	1
	$B^0 \to (\phi \to K^+K^-)\pi^+\pi^-$	8.8×10^{-8}	~ 0	~ 0	~ 0
	$B^0 \to p\overline{p}\pi^+\pi^-$	2.9×10^{-6}	~ 0	~ 0	~ 0
	$B_s^0 \to K^{*0}(\to K^+\pi^-)\rho^0$	$< 7.7 \times 10^{-4}$	~ 0	~ 0	~ 0
	$B_s^0 \to K^{*0}(\to K^+\pi^-)\overline{K}^{*0}(\to K^-\pi^+)$	1.1×10^{-6}	~ 0	~ 0	~ 0
	$B_s^0 \to p\overline{p}\pi^+\pi^-$	4.3×10^{-7}	~ 0	~ 0	~ 0
	$B_s^0 \to (\phi \to K^+K^-)\pi^+\pi^-$	1.7×10^{-6}	~ 0	~ 0	~ 0
	$B_s^0 \to \rho^0\rho^0$	$< 3.2 \times 10^{-4}$	~ 0	~ 0	~ 0
Partially reconstructed	$B_s^0 \to \eta'(\to \rho^0\gamma)J/\psi$	5.9×10^{-6}	340	400	890
	$B_s^0 \to J/\psi\eta'(\to \pi^+\pi^-\pi^0)$	7.8×10^{-8}	~ 0	1	2
	$B_s^0 \to J/\psi\eta' \to (\omega(\to \pi^+\pi^-)\gamma)$	8.3×10^{-9}	~ 0	~ 0	1
	$B_s^0 \to J/\psi\phi(\to \pi^+\pi^-\pi^0)$	5.6×10^{-7}	5	6	13
	$B_s^0 \to J/\psi\eta(\to \pi^+\pi^-\gamma)$	9.6×10^{-7}	7	8	18
	$B_s^0 \to \psi(2S)\eta(\to \pi^+\pi^-\gamma)$	1.0×10^{-7}	~ 0	~ 0	~ 0
	$B_s^0 \to J/\psi\eta(\to \pi^+\pi^-\pi^0)$	5.5×10^{-6}	~ 0	~ 0	~ 0
	$B^0 \to J/\psi\omega(\to \pi^+\pi^-\pi^0)$	1.2×10^{-6}	~ 0	~ 0	~ 0
	$B^0 \to J/\psi\eta(\to \pi^+\pi^-\pi^0)$	1.7×10^{-7}	~ 0	~ 0	~ 0
	$B^0 \to J/\psi\eta'(\to \pi^+\pi^-\pi^0)$	1.7×10^{-9}	~ 0	~ 0	~ 0
	$B^0 \to \psi(2S)(\to \eta(\to \gamma\gamma)J/\psi)\pi^+\pi^-$	1.7×10^{-5}	~ 0	~ 0	~ 0
	$B^+ \to J/\psi K^+\pi^-\pi^+$	6.5×10^{-6}	~ 0	~ 0	~ 0
	$B^+ \to \psi(2S)K^+\pi^-\pi^+$	1.3×10^{-5}	~ 0	~ 0	~ 0
	$B^+ \to J/\psi K^+K^-\pi^+$	3.6×10^{-6}	~ 0	~ 0	~ 0
	$B^+ \to J/\psi\omega(\to \pi^+\pi^-\pi^0)K^+$	1.7×10^{-5}	~ 0	~ 0	~ 0
	$B_s^0 \to \eta'(\to \rho^0\gamma)\mu^+\mu^-$	7.5×10^{-8}	3	3	7
	$B_s^0 \to \eta(\to \pi^+\pi^-\gamma)\mu^+\mu^-$	1.4×10^{-9}	~ 0	~ 0	~ 0
	$B_s^0 \to \eta(\to \pi^+\pi^-\pi^0)\mu^+\mu^-$	7.2×10^{-9}	~ 0	~ 0	~ 0
	$B_s^0 \to D_s^-(\to \phi(\to \pi^+\pi^-)\mu^+\overline{\nu}_\mu)\mu^-\nu_\mu$	1.2×10^{-7}	~ 0	~ 0	~ 0
	$B_s^0 \to \eta'(\to \omega(\to \pi^+\pi^-)\gamma)\mu^+\mu^-$	1.1×10^{-10}	~ 0	~ 0	~ 0
	$B_s^0 \to (D_s^-(\to \eta(\to \pi^+\pi^-\pi^0)\mu^+\overline{\nu}_\mu)\mu^-\nu_\mu$	1.1×10^{-7}	~ 0	~ 0	~ 0
	$B_s^0 \to (D_s^-(\to K^-\eta(\to \mu^+\mu^-\gamma))\pi^+$	1.7×10^{-9}	~ 0	~ 0	~ 0
	$B_s^0 \to \eta'(\to \pi^+\pi^-\eta(\to \gamma\gamma))\mu^+\mu^-$	4.4×10^{-8}	~ 0	~ 0	~ 0
	$B_s^0 \to \phi(\to \pi^+\pi^-\pi^0)\mu^+\mu^-$	7.1×10^{-9}	~ 0	~ 0	~ 0
	$B^0 \to D^-(\to \rho^0\mu^-\overline{\nu}_\mu)\mu^+\nu_\mu$	5.3×10^{-5}	~ 0	~ 0	~ 0
	$B^0 \to D^0(\to \pi^+\mu^-\overline{\nu}_\mu)\pi^-\mu^+\nu_\mu$	1.0×10^{-5}	~ 0	~ 0	~ 0
	$B^0 \to D^-(\to (\eta\mu^+\mu^-\gamma)\pi^-)\pi^+$	2.9×10^{-9}	~ 0	~ 0	~ 0
	$B^0 \to D^-(\to (\eta'(\to \mu^+\mu^-\gamma)\pi^-)\pi^+$	3.9×10^{-9}	~ 0	~ 0	~ 0
	$B^0 \to D^0(\to K^+\mu^-\overline{\nu}_\mu)\pi^-\mu^+\nu_\mu$	1.4×10^{-4}	~ 0	~ 0	~ 0
Over reconstructed	$B^+ \to J/\psi\pi^+$	2.5×10^{-6}	21	15	29
	$B^+ \to J/\psi K^+$	6.0×10^{-5}	9	3	8
	$B^+ \to \psi(2S)\pi^+$	1.9×10^{-7}	~ 0	~ 0	~ 0
	$B^+ \to \psi(2S)K^+$	5.0×10^{-6}	~ 0	~ 0	~ 0
	$B_c^+ \to J/\psi\pi^+$	6.8×10^{-6}	~ 0	~ 0	~ 0
	$B_c^+ \to J/\psi K^+$	5.4×10^{-7}	~ 0	~ 0	~ 0
	$B^+ \to \pi^+\mu^+\mu^-$	1.8×10^{-8}	~ 0	~ 0	~ 0
	$B^+ \to K^+\mu^+\mu^-$	4.4×10^{-7}	~ 0	~ 0	~ 0

Table 5.5: Classes of different peaking backgrounds considered within the analysis with branching fractions, B, taken from Ref. [175] or estimated using similar decays. For each background the yields are estimated for each of three datasets separately.

Table 5.6: The full list of selection requirements applied to select the candidates used within the control mode, resonant mode and rare mode invariant mass fits. The requirements here are made in addition to the stripping, trigger and PID fiducial requirements.

$B^0 \to J/\psi K^{*0}$ Control Mode	$B^0(B_s^0) \to J/\psi \pi^+\pi^-$ Resonant Modes	$B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ Rare Modes
BDT > 0.6	BDT > 0.6	BDT > 0.6
$m(K^+\pi^-) \in [822, 962]\,\mathrm{MeV}/c^2$	$m(\pi^+\pi^-) \in [515, 1300]\,\mathrm{MeV}/c^2$	$m(\pi^+\pi^-) \in [515, 1300]\,\mathrm{MeV}/c^2$
$m(K^+\pi^-\mu^+\mu^-) \in [5170, 6370]\,\mathrm{MeV}/c^2$	$m(\pi^+\pi^-\mu^+\mu^-) \in [5170, 6370]\,\mathrm{MeV}/c^2$	$m(\pi^+\pi^-\mu^+\mu^-) \in [5170, 5800]\,\mathrm{MeV}/c^2$
$m(K^+\mu^-) \notin [3076, 3116]\,\mathrm{MeV}/c^2$	$m(\pi^+\mu^-) \notin [3076, 3116]\,\mathrm{MeV}/c^2$	$m(\pi^+\mu^-) \notin [3076, 3116]\,\mathrm{MeV}/c^2$
$m(K^-\mu^+) \notin [3076, 3116]\,\mathrm{MeV}/c^2$	$m(\pi^-\mu^+) \notin [3076, 3116]\,\mathrm{MeV}/c^2$	$m(\pi^-\mu^+) \notin [3076, 3116]\,\mathrm{MeV}/c^2$
Opening Angle > 0.5 mrad	Opening Angle > 0.5 mrad	Opening Angle > 0.5 mrad
K^+!IsMuon	π^+!IsMuon	π^+!IsMuon
π^-!IsMuon	π^-!IsMuon	π^-!IsMuon
K^+ ProbNNK *(1-ProbNNπ)*(1-ProbNNp) > 0.2	π^+ ProbNNπ *(1-ProbNNK)*(1-ProbNNp) > 0.6	π^+ ProbNNπ *(1-ProbNNK)*(1-ProbNNp) > 0.6
π^- ProbNNπ *(1-ProbNNK)*(1-ProbNNp) > 0.2	π^- ProbNNπ *(1-ProbNNK)*(1-ProbNNp) > 0.2	π^- ProbNNπ *(1-ProbNNK)*(1-ProbNNp) > 0.2
μ^+ ProbNN$\mu > 0.1$	μ^+ ProbNN$\mu > 0.1$	μ^+ ProbNN$\mu > 0.1$
μ^- ProbNN$\mu > 0.1$	μ^- ProbNN$\mu > 0.1$	μ^- ProbNN$\mu > 0.1$

CHAPTER 6

Decay Model

“ The role of the infinitely small in nature is infinitely great. ”

Louis Pasteur,

As has been mentioned throughout this thesis, the accuracy of the final branching fraction measurement relies on an accurate estimation of sets of efficiencies that are derived from MC simulation samples. In Chapter 5 some of the drawbacks of the LHCb simulation were discussed, where the MC fails to accurately model certain kinematic and PID variables. The analysis then corrects for these inaccuracies by weighting the simulation samples via data-driven methods. This chapter is dedicated to another correction, which is made to the MC samples in order to further improve the data-MC agreement. This correction addresses the fact that the standard LHCb MC samples only consider a single resonance contribution to the final state hadron pair. For example, the $B^0 \to \rho^0 \mu^+ \mu^-$ MC sample, which is used as a proxy for the $B^0 \to \pi^+ \pi^- \mu^+ \mu^-$ mode only considers the decays via a $\rho^0(770)$ resonance to the dipion final state. In reality it is known that there are additional contributions towards the dipion final state, such as the $\rho^0(1410)$. Although these additional resonant contributions are small, their effect on some of the calculated efficiencies can be large enough to affect the final result.

In order to include the effect of these other resonance contributions when calculating efficiencies, a weighting procedure is used whereby the original MC decay model used to generate the samples is effectively replaced with a new model. To achieve this, the i^{th} candidate receives a

weight given by

$$\alpha_i = \frac{1}{\alpha_{\rm norm}} \frac{f_i'(m_{hh}, m_{\mu\mu}, \vec{\Omega})}{f_i(m_{hh}, m_{\mu\mu}, \vec{\Omega})}, \tag{6.1}$$

where $f(m_{hh}, m_{\mu\mu}, \vec{\Omega})$ and $f'(m_{hh}, m_{\mu\mu}, \vec{\Omega})$ refer to the original and new decay model. The angular distribution of the candidate is represented by $\vec{\Omega}$. Finally, the $\alpha_{\rm norm}$ term is a normalisation factor, which satisfies

$$\frac{1}{N} \sum_{i=0}^{N} \frac{1}{\alpha_{\rm norm}} \frac{f_i'(m_{hh}, m_{\mu\mu}, \vec{\Omega})}{f_i(m_{hh}, m_{\mu\mu}, \vec{\Omega})} = 1, \tag{6.2}$$

where N is the total number of candidates in the sample.

6.1 Angular Basis

As has been done in several previous analyses, it is possible to parameterise a given pseudoscalar B meson decay in terms of angular variables [155], [178]. An angular basis is normally used which describes the final state in terms of three separate helicity angles, visualised in Figure 6.1. The definition of the three helicity angles slightly varies depending upon whether the final state is a CP eigenstate, i.e. contains a $\pi^+\pi^-$ pair, or is not a CP eigenstate, e.g. contains a $K^+\pi^-$ pair. The decay angles are defined as follows:

- θ_h is the angle between the direction of the π^+ in the $\pi^+\pi^-$ rest-frame and the direction of the $\pi^+\pi^-$ system in the rest-frame of the B meson. In the case of a $K^+\pi^-$ final state the equivalent angle between the K^+ in the $K^+\pi^-$ frame and the direction of the $K^+\pi^-$ in the rest frame of the B meson, regardless of the charge of the kaon.

- θ_l is the angle between the direction of the μ^+ in the dimuon rest-frame and the direction of the $\mu^+\mu^-$ pair in the b-hadron rest-frame for a $\pi^+\pi^-$ final state. If instead the final state contains a $K^+\pi^-$ pair, the charge of the muon used to calculate the angle is the same as the charge of the kaon.

- ϕ is the angle between the two decay planes each of which is defined using a perpendicular vector. The first decay plane is defined as containing the $\pi^+\pi^-$ ($K^+\pi^-$) system and the other defined as containing the $\mu^+\mu^-$ system. In the $\pi^+\pi^-$ case the plane vector is defined with a positive pion track crossed with a negative pion track. For the $K^+\pi^-$ case, the hadron plane is defined by the cross-product of the momentum of the kaon and the momentum of the pion in the rest frame of the B meson. The dimuon plane is then defined by the cross product of μ_1 and μ_2 where μ_1 shares the same charge as the kaon.

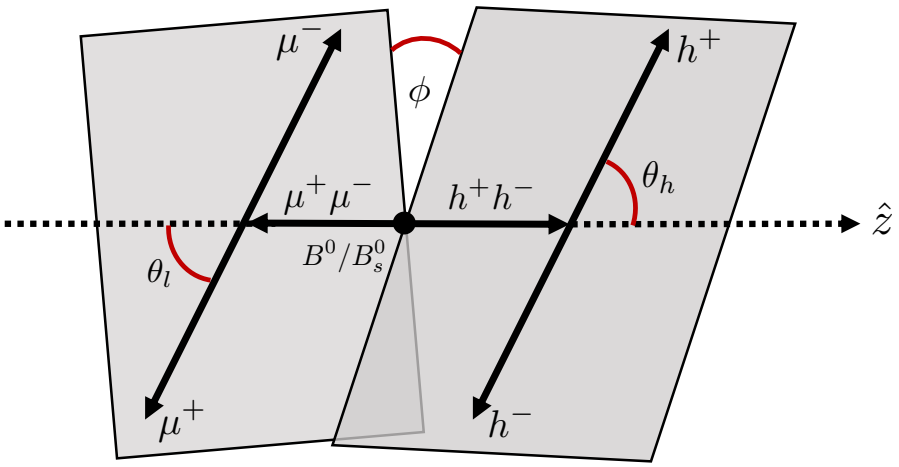

Figure 6.1: A visual representation of the three helicity angles used in the decay model angular basis, where h refers to a final state hadron.

6.2 Resonant Mode Decay Model

For the resonant mode analysis the $B^0 \to J/\psi K^+\pi^-$, $B^0 \to J/\psi \pi^+\pi^-$ and $B_s^0 \to J/\psi \pi^+\pi^-$ MC samples are all corrected to include additional resonant contributions. The resonant mode decay model is introduced by considering the time dependent decay rate for $B^0 \to J/\psi \pi^+\pi^-$ which is described in detail in [179]. The differential decay rate is written as

$$\begin{aligned}\frac{\mathrm{d}^5\Gamma}{\mathrm{d}t\,\mathrm{d}m_{hh}\,\mathrm{d}\theta_l\,\mathrm{d}\theta_h\,\mathrm{d}\phi} = Ne^{-\Gamma t}\Bigg(&\frac{|A|^2 + \left|(q_{\mathrm{os}}/p_{\mathrm{os}})\overline{A}\right|^2}{2}\cosh\left(\frac{\Delta\Gamma t}{2}\right)\\ &+ \frac{|A|^2 - \left|(q_{\mathrm{os}}/p_{\mathrm{os}})\overline{A}\right|^2}{2}\cos\left(\Delta m t\right)\\ &- \mathcal{R}e\left(\frac{q_{\mathrm{os}}}{p_{\mathrm{os}}}\right)A^*\overline{A}\Big)\sinh\left(\frac{\Delta\Gamma t}{2}\right)\\ &- \mathcal{I}m\left(\frac{q_{\mathrm{os}}}{p_{\mathrm{os}}}\right)A^*\overline{A}\Big)\sin(\Delta m t)\Bigg).\end{aligned} \tag{6.3}$$

The equivalent decay rate for the production conjugate mode is given by

$$\begin{aligned}\frac{\mathrm{d}^5\overline{\Gamma}}{\mathrm{d}t\,\mathrm{d}m_{hh}\,\mathrm{d}\theta_l\,\mathrm{d}\theta_h\,\mathrm{d}\phi} = \left|\frac{p_{\rm os}}{q_{\rm os}}\right|^2 Ne^{-\Gamma t}\Bigg(&\frac{|A|^2+\left|(q_{\rm os}/p_{\rm os})\overline{A}\right|^2}{2}\cosh\Big(\frac{\Delta\Gamma t}{2}\Big)\\ &-\frac{|A|^2-\left|(q_{\rm os}/p_{\rm os})\overline{A}\right|^2}{2}\cos\Big(\Delta m t\Big)\\ &-\mathcal{R}e\Big(\frac{q_{\rm os}}{p_{\rm os}}\Big)A^*\overline{A}\Big)\sinh\Big(\frac{\Delta\Gamma t}{2}\Big)\\ &+\mathcal{I}m\Big(\frac{q_{\rm os}}{p_{\rm os}}\Big)A^*\overline{A}\Big)\sin(\Delta m t)\Bigg).\end{aligned} \tag{6.4}$$

In Equation 6.3 and Equation 6.4 N is a constant; A and $\overline{A}$ are functions of m_{hh} and the helicity angles and are the amplitude of $B^0 \to J/\psi\pi^+\pi^-$ and $\overline{B}^0 \to J/\psi\pi^+\pi^-$ decays respectively at decay time $t = 0$; Δm is the difference between the heavy and light B mass eigenstates; $\Delta\Gamma$ is the width difference and $p_{\rm os}$ and $q_{\rm os}$ are the complex parameters describing the relationship between the mass and flavour eigenstates of the B meson oscillations. By assuming that $|p_{\rm os}/q_{\rm os}|$ is equal to unity, summing together the contributions from the two processes and integrating out the time dependence, the total decay rate, $S(m_{hh},\theta_h,\theta_l,\phi)$, can be expressed as

$$\begin{aligned}S(m_{hh},\theta_h,\theta_l,\phi) = &\left|A(m_{hh},\theta_h,\theta_l,\phi)\right|^2 + \left|\overline{A}(m_{hh},\theta_h,\theta_l,\phi)\right|^2\\ &-2\mathcal{D}\mathcal{R}e\Big(\frac{q_{\rm os}}{p_{\rm os}}A^*(m_{hh},\theta_h,\theta_l,\phi)\overline{A}(m_{hh},\theta_h,\theta_l,\phi)\Big)\end{aligned} \tag{6.5}$$

where $\mathcal{D}$ is a term which accounts for both time dependent detection efficiencies and the width difference between the light and heavy B meson mass eigenstates. The total decay rate can be rewritten to explicitly show the summation over the different resonance contributions. Here, the decay rate is expressed in terms of a summation over the final state dilepton helicity, λ_ψ, defined as the difference between the two final state muon helicities, as is done in [180]. The decay rate is written as

$$\begin{aligned}S(m_{hh},\vec{\Omega}) = p(m_{hh})q(m_{hh})\sum_{\lambda_\psi=+1,0,-1}\Bigg(&\left|\sum_{\rm res}A_{{\rm res},\lambda_\psi}(m_{hh},\vec{\Omega})\right|^2 + \left|\sum_{\rm res}\overline{A}_{{\rm res},\lambda_\psi}(m_{hh},\vec{\Omega})\right|^2\\ &-2\mathcal{D}\Big(\frac{q_{\rm os}}{p_{\rm os}}A^*_{{\rm res},\lambda_\psi}(m_{hh},\vec{\Omega})\overline{A}_{{\rm res},\lambda_\psi}(m_{hh},\vec{\Omega})\Big)\Bigg)\end{aligned} \tag{6.6}$$

where $\vec{\Omega}$ represents the helicity angle information, $A_{\mathrm{res},\lambda_\psi}$ are complex helicity amplitudes which contain the physical dynamics of each resonance contribution and $p(m_{hh})$ and $q(m_{hh})$ are the daughter momenta of the hadron pair in the b hadron and h^+h^- rest frames, respectively.

By assuming an isobar model where a complex amplitude can be expressed as a coherent sum, $A_{\mathrm{res},\lambda_\psi}$, can be expressed as

$$A_{\mathrm{res},\lambda_\psi} = R(m_{hh}|m_{\mathrm{res}},\Gamma_{\mathrm{res}}) \sum_{\lambda_J} H_{\mathrm{res},\lambda_J} d^1_{\lambda_J,\lambda_\psi}(\theta_\ell) d^J_{\lambda_J,0}(\theta_h) e^{-i\lambda_J\phi} \,, \tag{6.7}$$

where J is the total angular momentum of the resonance, λ_J is the meson helicity, $d^J_{\lambda_1,\lambda_2}$ are Wigner D-functions [181] , $H_{\mathrm{res},\lambda_J}$ are complex helicity amplitudes and $R(m_{hh}|m_{\mathrm{res}},\Gamma_{\mathrm{res}})$ is a relativistic lineshape. For most resonances considered this is a relativistic Breit-Wigner function [182]. For certain resonances a Flatté function which is a coupled, two-channel Breit-Wigner function is particularly useful to describe broad lineshapes near thresholds [183].

For decays to a $K^+\pi^-$ state the decay rate is given by

$$S(m_{hh},\vec{\Omega}) = p(m_{hh})q(m_{hh}) \sum_{\lambda_\psi=+1,0,-1} \left(\left| \sum_{\mathrm{res}} A_{\mathrm{res},\lambda_\psi}(m_{hh},\vec{\Omega}) \right|^2 + \left| \sum_{\mathrm{z}} A_{z,\lambda_\psi}(m_{hh},\vec{\Omega}) \right|^2 \right) \tag{6.8}$$

which is based on the model previously used by the Belle collaboration to describe the $J/\psi\, K^+\pi^-$ final state [184]. The additional term, $A_{\mathrm{z},\lambda_\psi}$, is an additional complex amplitude, which accounts for the contribution of exotic resonances decaying to $J/\psi\, K^+\pi^-$ which are treated with a separate amplitude following the method seen in [23]. As $K^+\pi^-$ is odd under a CP transformation, no conjugate process or interference term is included in the decay rate.

Each of the three weighted simulation samples include different resonance contributions and the list of resonances included, along with the corresponding helicity amplitude values, are taken from previous analyses. The helicity amplitude information is given in Table 6.1.

Resonance	$\lvert H_0\rvert$	$\arg(H_0)$	$\lvert H_+\rvert$	$\arg(H_+)$	$\lvert H_-\rvert$	$\arg(H_-)$
$B^0 \to J/\psi K^+\pi^-$						
$K_0^*(800)$	1.12 ± 0.04	2.30 ± 0.04	—	—	—	—
$K^*(892)$	1.0 (fixed)	0.0 (fixed)	$(8.44 \pm 0.10) \times 10^{-1}$	3.14 ± 0.03	$(1.96 \pm 0.14) \times 10^{-1}$	-1.70 ± 0.07
$K^*(1410)$	$(1.19 \pm 0.27) \times 10^{-1}$	0.81 ± 0.26	$(1.23 \pm 0.38) \times 10^{-1}$	-1.04 ± 0.26	$(0.36 \pm 0.39) \times 10^{-1}$	0.67 ± 1.06
$K_0^*(1430)$	$(8.90 \pm 0.28) \times 10^{-1}$	-2.17 ± 0.05	—	—	—	—
$K_2^*(1430)$	4.66 ± 0.18	-0.32 ± 0.05	4.65 ± 0.18	-3.05 ± 0.08	1.26 ± 0.23	-1.92 ± 0.20
$K^*(1680)$	$(1.39 \pm 0.43) \times 10^{-1}$	-2.46 ± 0.31	$(0.82 \pm 0.48) \times 10^{-1}$	-2.85 ± 0.49	$(1.61 \pm 0.56) \times 10^{-1}$	1.88 ± 0.28
$K_3^*(1780)$	16.8 ± 3.6	-1.43 ± 0.24	19.1 ± 4.5	2.03 ± 0.31	10.2 ± 5.2	1.55 ± 0.62
$K_0^*(1950)$	$(2.41 \pm 0.60) \times 10^{-1}$	-2.39 ± 0.25	—	—	—	—
$K_2^*(1980)$	4.53 ± 0.74	-0.26 ± 0.16	3.78 ± 0.98	3.08 ± 0.28	3.51 ± 1.03	2.63 ± 0.34
$K_4^*(2045)$	590 ± 136	-2.66 ± 0.23	676 ± 164	0.06 ± 0.25	103 ± 174	-1.03 ± 1.62
$Z^+(4430)$	1.12 ± 0.32	-0.31 ± 0.26	1.17 ± 0.46	0.77 ± 0.25	1.17 ± 0.46	0.77 ± 0.25
$Z^+(4200)$	0.71 ± 0.37	2.14 ± 0.40	3.23 ± 0.79	3.00 ± 0.15	3.23 ± 0.79	3.00 ± 0.15
$B^0 \to J/\psi \pi^+\pi^-$						
$\rho(770)$	0.758	0.000	-0.027	-0.057	-0.647	-0.057
$f_0(500)$	0.451	-0.192	—	—	—	—
$f_2(1270)$	-1.637	-0.955	-0.326	-0.017	1.337	-0.463
$w(782)$	-0.006	-0.261	-0.147	0.169	-0.107	0.212
$\rho(1450)$	-0.159	-0.047	-0.039	0.075	0.080	0.080
$\rho(1700)$	0.036	0.106	0.031	0.073	0.091	-0.065
$B_s^0 \to J/\psi \pi^+\pi^-$						
$f_0(980)$	1.00	0.00	—	—	—	—
$f_0(1500)$	7.35	2.41	—	—	—	—
$f_2(1270)$	10.91	1.68	13.77	-0.74	13.76	-2.40
$f_0(1790)$	2.82	1.36	—	—	—	—
$f_2'(1525)$	18.88	-2.30	13.14	2.35	7.17	-0.06

Table 6.1: Model parameters taken from the Belle collaboration fit to $B^0 \to J/\psi K^+\pi^-$ as described in [184], and from the LHCb fits to $B^0 \to J/\psi \pi^+\pi^-$ as described in [180] and $B_s^0 \to J/\psi \pi^+\pi^-$ as described in [173]. The uncertainties on the $B^0 \to J/\psi K^+\pi^-$ parameters are purely statistical in nature, however, systematic uncertainties are small. The amplitudes for the $J/\psi \pi^+\pi^-$ resonances are converted from the fit fraction information given within the papers. All of the resonances are modelled with a relativistic Breit-Wigner function apart from the $f_0(980)$ which is modelled with a Flatté function.

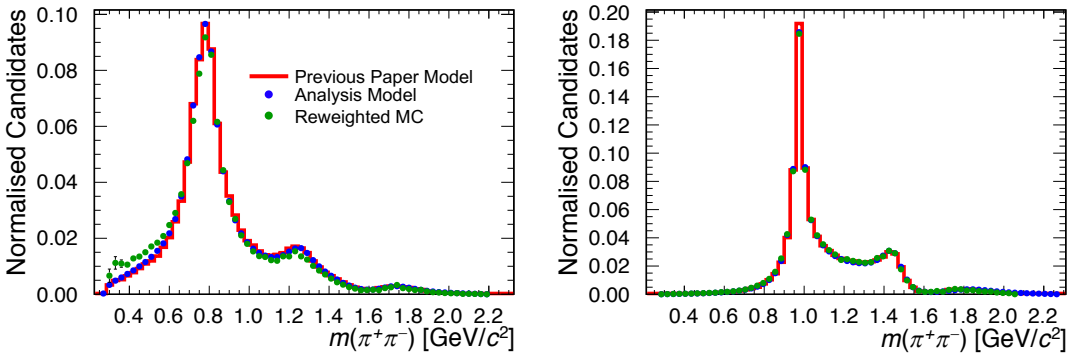

Figure 6.2: Left: Comparison between the decay model described in [180], the output of the decay model from this analysis and the result of weighting the $B^0 \to J/\psi\,\pi^+\pi^-$ MC. Right: The equivalent comparison between the decay model described in [173], the decay model output from this analysis and the result of weighting the $B^0_s \to J/\psi\,\pi^+\pi^-$ MC.

The decay model requires mass, spin and width information for each resonant contribution. The three MC resonant samples use the parameters shown in Table 6.3.

The decay model for the $B^0 \to J/\psi\,\pi^+\pi^-$ and $B^0_s \to J/\psi\,\pi^+\pi^-$ modes can be compared to the results from the respective LHCb analyses. This is done to check that the implementation of the decay models within software is consistent with the original models. Figure 6.2 shows this comparison in addition to the effect of applying the decay model weights on the MC samples. The left-hand figure of Figure 6.2 shows the comparison for the B^0 mode and the right the equivalent for the B^0_s mode. As can be seen, the implementation of the decay model within this analysis agrees well with the models from the previous analyses. The final weighted MC distributions also show a good agreement with these models. This check is not made for the $B^0 \to J/\psi\,K^+\pi^-$ decay model implemented from the Belle analysis as the equivalent output of the original model is not available for comparison.

After the application of the decay model, the dihadron mass distributions from the three MC samples, with and without the decay model weights applied, can be compared to one another. As can be seen from Figure 6.3 the mass distributions are altered after the application of the weights such that the inclusion of the decay model resonances can be clearly seen.

In addition, as a check of the decay mode weighting, the $m(\pi^+\pi^-)$ distribution of the $B^0 \to J/\psi\,\pi^+\pi^-$ and $B^0_s \to J/\psi\,\pi^+\pi^-$ modes after the completion of the resonant mode fits of Chapter 8 can be compared to the weighted MC distributions. In order to extract the data $m(\pi^+\pi^-)$ distribution, the *sPlot* [166] method is used to unfold the signal distribution by applying a series of *sPlot* weights, similarly as was done within Chapter 5. Figure 6.4 shows this comparison and shows that for the resonant mode, the weighted MC distributions agree well with

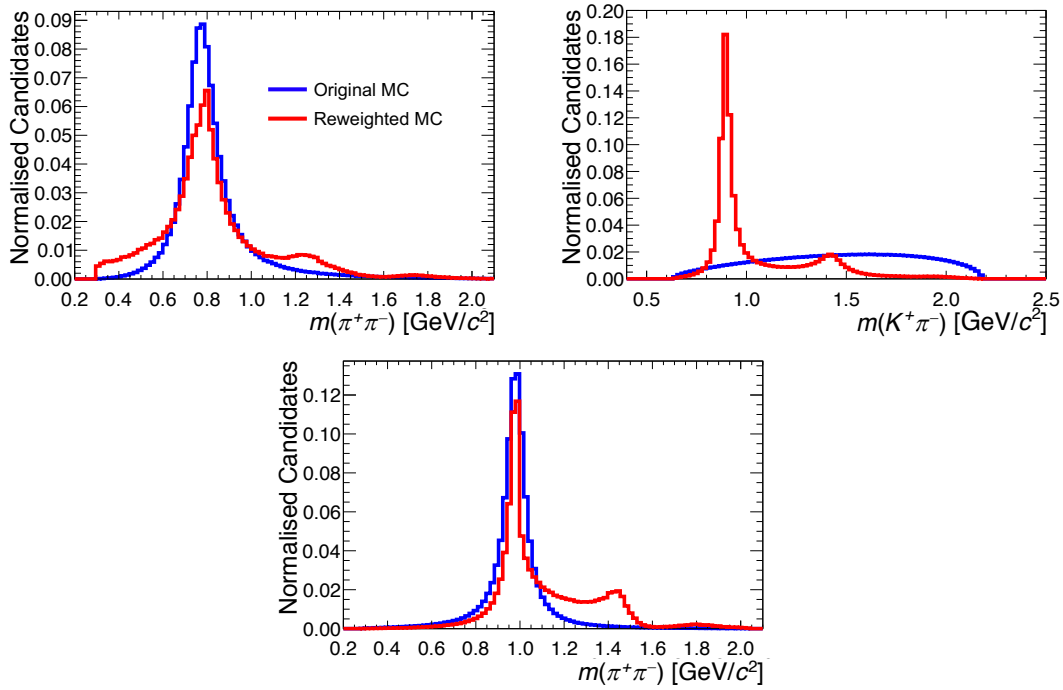

Figure 6.3: The effect of the application of the decay model weighting on the $B^0 \to J/\psi\pi^+\pi^-$ (top left), $B^0 \to J/\psi K^+\pi^-$ (top right) and $B_s^0 \to J/\psi\pi^+\pi^-$ (bottom) samples. The inclusion of the additional resonances on the dihadron mass after the application of the weights can be clearly seen.

the unfolded data distributions.

6.3 Rare Mode Decay Model

A similar procedure is followed to weight the rare mode MC samples for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$, $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ decays to include the contribution of additional resonances. Unlike the resonant mode weighting however, the rare mode decays have a dimuon pair which do not originate from a J/ψ meson. As a result, the decay model built for the rare mode must also include the variation of the decay rate as a function of q^2 which means that a different approach is required. The general formulation of the rare mode decay model is taken from previous theoretical studies such as described in [185].

The rare mode decay model uses the full angular distribution which, in it's simplest form

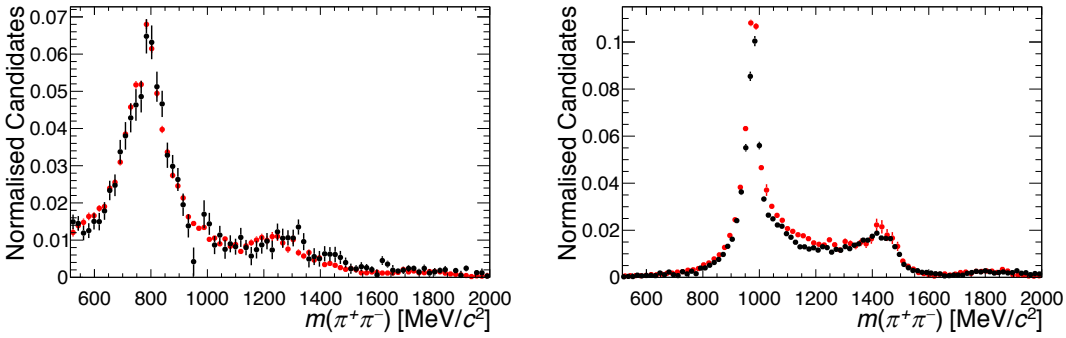

Figure 6.4: Comparison between the resonant mode MC samples weighted with the full decay model and the $\pi^+\pi^-$ distributions extracted from the resonant mode fits of Chapter 8 for the 2017+2018 dataset. The left plot shows the $B^0 \to J/\psi\,\pi^+\pi^-$ comparison and the right plot the $B^0_s \to J/\psi\,\pi^+\pi^-$ comparison. Similar plots are seen for the other datasets.

can be written as a function of q^2 and the decay angles as

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos(\theta_l)\,\mathrm{d}\cos(\theta_h)\,\mathrm{d}\phi} = \frac{3}{32\pi} I(q^2, \theta_l, \theta_h, \phi). \tag{6.9}$$

Equation 6.9 can then be expanded in terms of different angular coefficients. This expansion takes the form

$$\begin{aligned} I(q^2, \theta_l, \theta_h, \phi) = {} & I_1^s \sin^2\theta_h + I_1^c \cos^2\theta_h + (I_2^s \sin^2\theta_h + I_2^c \cos^2\theta_h)\cos 2\theta_l \\ & + I_3 \sin^2\theta_h \sin^2\theta_l \cos 2\phi + I_4 \sin 2\theta_h \sin 2\theta_l \cos\phi \\ & + I_5 \sin 2\theta_h \sin\theta_l \cos\phi + I_6^s \sin^2\theta_h \cos\theta_l \\ & + I_7 \sin 2\theta_h \sin\theta_l \sin\phi \\ & + I_8 \sin 2\theta_h \sin 2\theta_l \sin\phi + I_9 \sin^2\theta_h \sin^2\theta_l \sin 2\phi. \end{aligned} \tag{6.10}$$

Each of the individual angular coefficients are themselves a physical observable and they can be expressed in terms of different complex polarisation amplitudes similarly to as is done for the resonant mode decay model. The individual angular coefficients are defined, for a P-wave, state by

$$I_1^s = \frac{2+\beta_\mu^2}{4}[|A_\perp^L| + |A_\parallel^L|^2 + (L \to R)] + (1-\beta_\mu^2)\mathcal{R}(A_\perp^L A_\perp^{R*} + A_\parallel^L A_\parallel^{R*}), \tag{6.11}$$

$$I_1^c = |A_0^L|^2 + |A_0^R|^2 + (1-\beta_\mu^2)(2\mathcal{R}(A_0^L A_0^R)), \tag{6.12}$$

$$I_2^s = \frac{\beta_\mu^2}{4}[|A_\perp^L| + |A_\parallel^L|^2 + (L \to R)], \tag{6.13}$$

$$I_2^c = -\beta_\mu^2[|A_0^L|^2 + (L \to R)], \tag{6.14}$$

$$I_3 = \frac{1}{2}\beta_\mu^2[|A_\perp^L| - |A_\parallel^L|^2 + (L \to R)], \tag{6.15}$$

$$I_4 = \frac{1}{\sqrt{2}}\beta_\mu^2[\mathcal{R}(A_0^L A_\parallel^{L*}) + (L \to R))], \tag{6.16}$$

$$I_5 = \sqrt{2}\beta_\mu\Big[\mathcal{R}(A_0^L A_\perp^{L*}) - (L \to R)\Big], \tag{6.17}$$

$$I_6^s = 2\beta_\mu[\mathcal{R}(A_\parallel^L A_\perp^{L*}) - (L \to R)], \tag{6.18}$$

$$I_7 = \sqrt{2}\beta_\mu\Big[\mathcal{I}(A_0^L A_\parallel^{L*}) - (L \to R)\Big], \tag{6.19}$$

$$I_8 = \frac{1}{\sqrt{2}}\beta_\mu^2[\mathcal{I}(A_0^L A_\perp^{L*}) + (L \to R)], \tag{6.20}$$

$$I_9 = \beta_\mu^2[\mathcal{I}(A_\parallel^{L*} A_\perp^L) + (L \to R)]. \tag{6.21}$$

where β_μ^2 is a kinematic term describing the lepton momentum in the dilepton rest frame which is given by

$$\beta_\mu^2 = 1 - \frac{4m_\mu^2}{q^2} \tag{6.22}$$

where m_μ is the muon mass and L and R refer to left and right-handed polarisation amplitudes. These terms are further modified when including the effect of an S-wave state and this is shown in Equation 3.1 of Ref. [186].

As was done for the resonant mode, a decay to a CP eigenstate requires a modification of the decay rate expression. This modification is shown in Equation 6.23

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos(\theta_l)\,\mathrm{d}\cos(\theta_h)\,\mathrm{d}\phi} = \frac{3}{32\pi}(I(q^2,\theta_l,\theta_h,\phi) + \bar{I}(q^2,\theta_l,\theta_h,\phi)) \tag{6.23}$$

which neglects the small interference term between the conjugate and non-conjugate processes and where the term with $I \to \bar{I}$ is equivalent to the sign of the $I_{5,6,8,9}$ terms flipping.

The transversity amplitudes $A_0^{L,R}$, $A_\parallel^{L,R}$, $A_\perp^{L,R}$ are related to the helicity amplitudes used within the resonant mode decay model via

$$\begin{aligned} A_0 &= H_0 \\ A_\perp &= \frac{1}{\sqrt{2}}(H_+ - H_-) \\ A_\parallel &= \frac{1}{\sqrt{2}}(H_+ + H_-). \end{aligned} \tag{6.24}$$

The transversity amplitudes are expressed in terms of form-factors relevant for $B^0 \to X\mu^+\mu^-$ processes, and Wilson Coefficients. The form-factors commonly defined within the theoretical literature, which are of relevance for a vector resonance, are $A_1, A_2, A_3, V, T_1, T_2, T_3$. The equivalent form-factors commonly defined for scalar resonances are f_T, f_+, f_0. These form-factors are often transformed to make them easier to analyse and evaluate with the additional form-factors, T_{23} and A_{12}, common examples which are defined as

$$A_{12} = \frac{(m_B + m_V)^2(m_B^2 - m_V^2 - q^2)A_1(q^2) - \lambda(q^2, m_B, m_V)A_2(q^2)}{16 m_B m_V^2 (m_B + m_V)} \tag{6.25}$$

and

$$T_{23} = \frac{(m_B^2 - m_V^2)(m_B^2 + 3m_V^2 - q^2)T_2(q^2) - \lambda(q^2, m_B, m_V)T_3(q^2)}{8 m_B m_V^2 (m_B - m_V)} \tag{6.26}$$

where λ is

$$\lambda(q^2, m_B, m_V) = (q^2)^2 + m_B^2 + m_V^2 - 2(m_V^2 m_B^2 + m_V^2 q^2 + m_B^2 q^2) \tag{6.27}$$

and where m_V refers to the pole mass of the resonance and m_B the mass of the mother particle. For a vector resonance the transversity amplitudes are

$$A_\parallel^{L,R} \propto (m_B^2 - m_{hh}^2)\Big\{\frac{(C_9^{\text{eff}} \mp C_{10})A_1(q^2)}{m_B - m_{hh}} + \frac{2C_7 m_b T_2(q^2)}{q^2}\Big\}, \tag{6.28}$$

$$A_\perp^{L,R} \propto \sqrt{\lambda_M}\Big\{\frac{(C_9^{\text{eff}} \mp C_{10})V(q^2)}{m_B + m_{hh}} + \frac{2C_7 m_b T_1(q^2)}{q^2}\Big\}, \tag{6.29}$$

C_1	C_2	C_3	C_4	C_5
-0.257	1.009	-0.005	-0.078	0.000
C_6	C_7	C_8	C_9	C_{10}
0.001	-0.304	-0.167	4.211	-4.103

Table 6.2: Values of the Standard Model Wilson coefficients which are used within the rare mode decay model weighting.

$$A_0^{L,R} \propto \frac{1}{2m_{hh}\sqrt{q^2}}\Big[(C_9^{\text{eff}} \mp C_{10})((m_B^2 - m_{hh}^2 - q^2)(m_B + m_{hh})A_1(q^2) - \frac{\lambda_M A_2(q^2)}{m_B + m_{hh}}) + 2m_b C_7((m_B^2 + 3m_{hh}^2 - q^2)T_2(q^2) - \frac{\lambda_M T_3(q^2)}{m_B^2 - m_{hh}^2})\Big], \tag{6.30}$$

where m_b is the mass of the bottom quark and m_B is the mass of the B^0 (B_s^0) meson.

For the scalar resonances both the parallel and the perpendicular amplitudes are zero by definition. Meanwhile the longitudinal component is given by

$$A_0^{L,R} \propto \frac{\lambda_M}{q^2}\Big\{(C_9^{\text{eff}} \mp C_{10})f_+(q^2) + \frac{2C_7 m_b f_T(q^2)}{m_B + m_{hh}}\Big\}. \tag{6.31}$$

Examples of the form-factor variations, as implemented within the analysis software, can be seen in Figure 6.5 which shows examples of both vector and scalar form-factors.

The decay model used assumes that the underlying physical behaviour is well described by the SM. As a result, the Wilson Coefficients used correspond to the current SM values which are listed in Table 6.2. The possibility of instead performing the rare mode weighting with an alternative set of Wilson Coefficients is evaluated as a source of systematic uncertainty as is discussed further in Chapter 9.

The different resonances which are included for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$, $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ decay model weighting are included in Table 6.3. References are provided for the different theoretical papers that the specific form-factor parameterisations for each resonance are taken from. The physical parameters, including the mass, width and spin are included in the table. Finally, when building the model, each resonance is modelled using a Breit-Wigner lineshape and the absolute contribution of each resonance to the decay rate is scaled with a coupling term. The coupling term is determined from branching and fit fraction information and replicates the relative contribution of each resonance seen in the corresponding resonant mode decays.

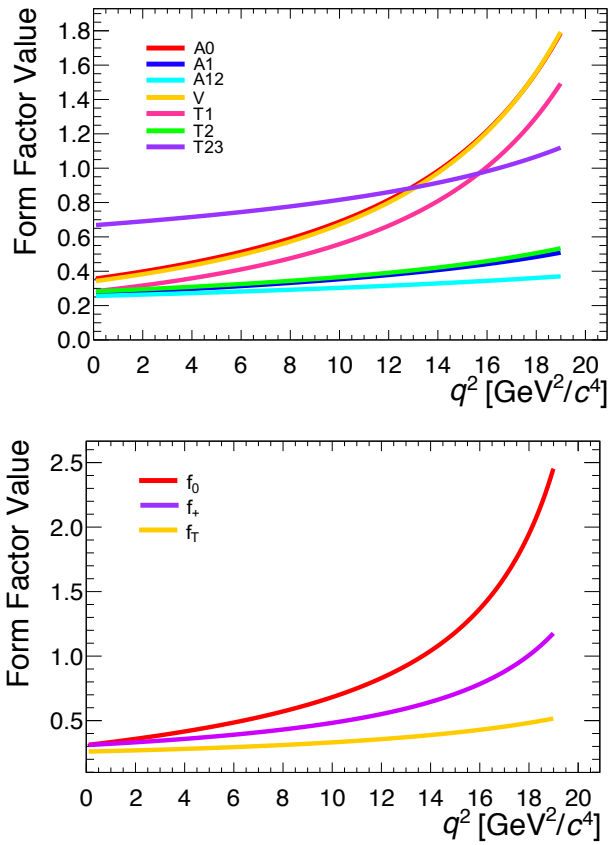

Figure 6.5: Examples of recreated vector form-factor parameters (top) and scalar form-factor parameters (bottom) used in the decay model as a function of q^2. The top plot shows the vector form-factors for the $K^*(892)$ resonance as described in [82]. The bottom plot shows the scalar form-factors for the $K_0^*(800)$ as described in [187].

The results of the decay model weighting for the rare mode MC samples can be seen in Figure 6.6, where the effect of including the additional resonance contributions can be clearly seen. As was mentioned previously, a possible source of systematic uncertainty resulting from the decay model weighting could arise from the choice of Wilson Coefficients used. Two other sources of systematic uncertainty, relating to the form-factor parameterisations used and the possibility of weighting with only a single resonance contribution are also considered. The single resonance weighting is motivated by the fact that the decay model for the rare mode is not yet well known experimentally however a single resonance is expected to dominate. This is discussed further in Chapter 9.

Given the desire to make measurements of the rare mode branching fractions in different regions of q^2 it is also useful to check the q^2 distributions seen in the rare mode MC samples

Resonance	Mass	Width	Spin	Coupling	Form Factor Paper
$B^0 \to \pi^+\pi^-\mu^+\mu^-$					
$\rho(770)$	0.775	0.148	1	0.693	[82]
$f_0(500)$	0.513	0.335	0	0.975	[188]
$w(782)$	0.783	0.0085	1	0.623	[82]
$f_2(1270)$	1.275	0.187	2	0.356	[189]
$\rho(1450)$	1.465	0.40	1	0.412	[88]
$\rho(1700)$	1.72	0.25	1	0.400	[88]
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$					
$f_0(980)$	0.990	0.070	0	0.859	[188]
$f_2(1270)$	1.270	0.185	2	0.072	[190]
$f_0(1500)$	1.4609	0.124	0	0.443	[188]
$f_2'(1525)$	1.522	0.084	2	0.081	[190]
$f_0(1790)$	1.790	0.270	0	0.271	[188]
$B^0 \to K^+\pi^-\mu^+\mu^-$					
$K^*(892)$	0.896	0.049	1	0.852	[82]
$K_0^*(800)$	0.824	0.478	0	1.000	[187]
$K_2^*(1430)$	1.426	0.0985	2	0.606	[190]
$K_0^*(1430)$	1.425	0.270	0	1.323	[187]
$K_1(1410)$	1.414	0.232	1	0.756	[88]

Table 6.3: Resonances included for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$, $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ decay model weighting.

before and after the weighting. This comparison can be seen in Figure 6.7

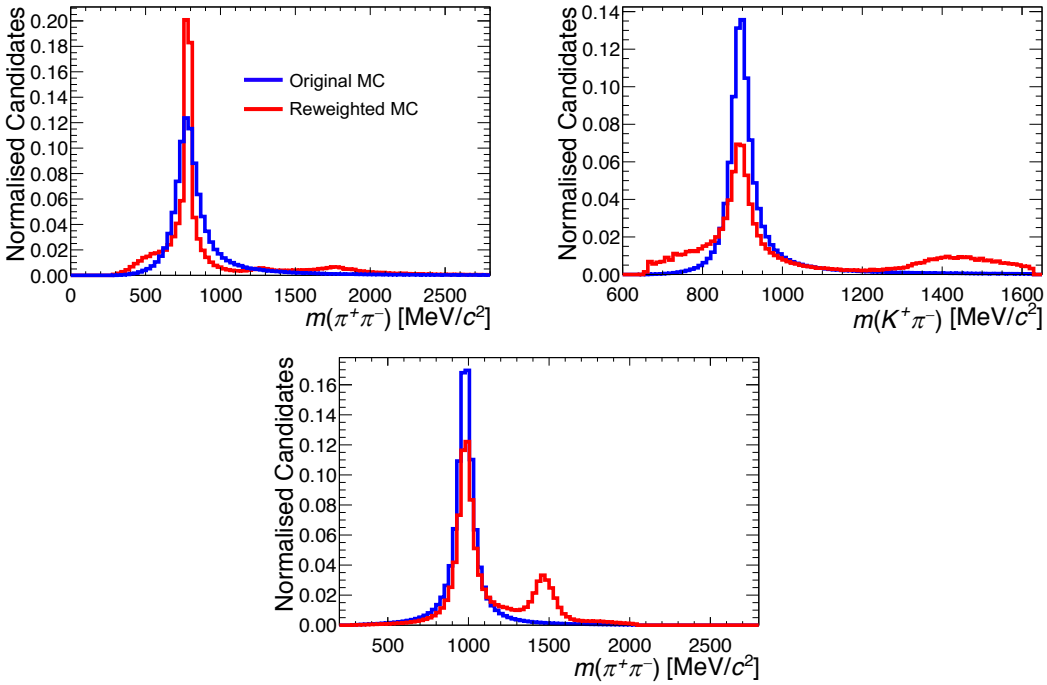

Figure 6.6: The result of the rare mode decay model weighting on the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC (top), the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC (middle) and the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ MC (bottom). The narrower width of the $\rho^0(770)$ resonance seen after the weighting is due to the different ρ^0 widths used in the LHCb MC software and within the weighting procedure.

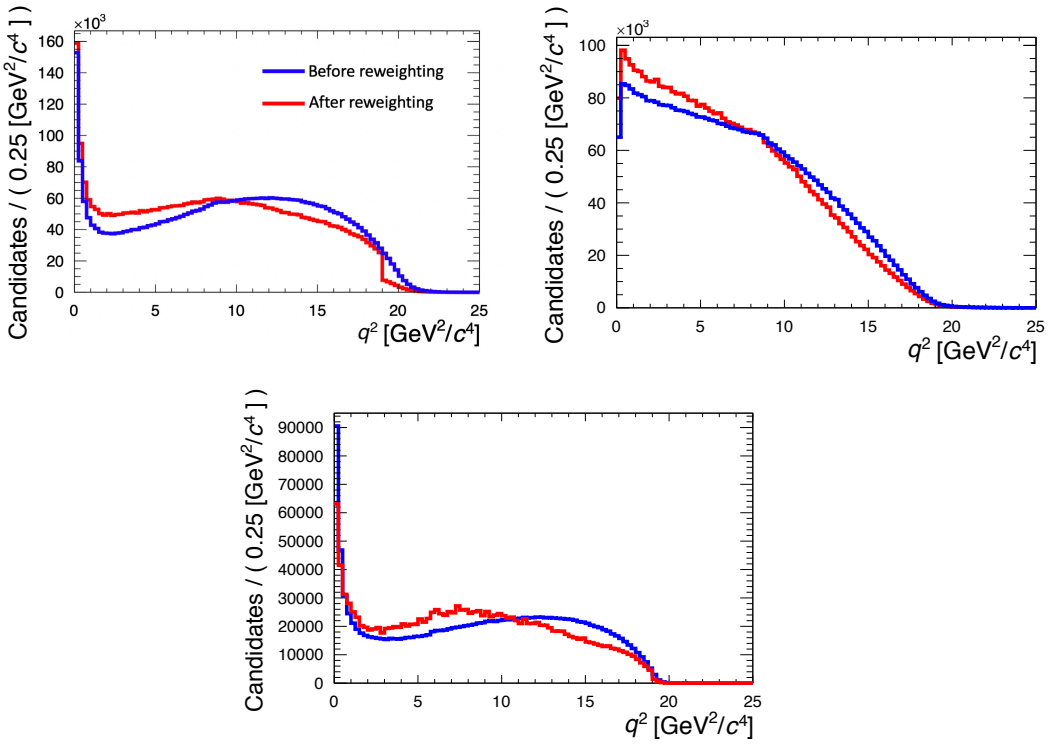

Figure 6.7: The q^2 distributions before and after the rare mode MC weighting. The top left plot shows the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC, the top right shows the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ MC and the bottom plot the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC. The discontinuity seen at 19 GeV^2/c^4 in the top left plot is a feature of the decay model software where candidates above 19 GeV^2/c^4 can lead to poorly defined form-factor values and these candidates are given a weight of zero.

CHAPTER 7

Efficiencies

“ Data! Data! Data! I can’t make bricks without clay! ”

Sherlock Holmes,

In order to provide an accurate measurement of the rare mode branching fractions it is necessary to determine a series of efficiencies that describe the fraction of signal and background candidates which survive each step of the selection procedure. Within this analysis efficiencies are calculated using simulated samples. In order to provide a branching fraction measurement in a particular bin of q^2, the full set of rare mode efficiencies are calculated using candidates which exclusively fall into that particular q^2 bin.

7.1 Definition, Types and Uncertainties

Within this thesis the efficiency which corresponds to a specific selection requirement is defined as the fraction of candidates within the sample after the application of the requirement, divided by the number of candidates within the sample before the requirement is made. Furthermore, each candidate is weighted by the product of the various correcting weights which are mentioned throughout this thesis. These corrections include the kinematic weighting described in Chapter 5, the decay model weighting of Chapter 6, PID corrections applied via the LHCb `PIDCalib` package and then individual corrections to the trigger efficiency, tracking efficiency and stripping-level PID requirements with these final three corrections all discussed within this chapter. The

efficiency definition is written as

$$\epsilon = \frac{\sum_{i=0}^{N_i} w_i}{\sum_{j=0}^{N_j} w_j} \tag{7.1}$$

where N_i is the total number of candidates in the dataset after the application of the requirement, N_j is the total number of candidates before the application of the requirement, and $w_{i(j)}$ are the product of the individual correcting weights. In order for this efficiency definition to remain meaningful, the correcting weights are applied to all of the MC samples used, including generator level and pre-stripping MC samples where appropriate.

The total efficiency for a decay mode, ϵ, can be written in terms of various sub-efficiencies as

$$\epsilon = \epsilon^{\texttt{Geo}} \times \epsilon^{\texttt{Sel}\cdot} \times \epsilon^{\texttt{PID}} \tag{7.2}$$

where $\epsilon^{\texttt{Geo}}$ is the geometrical acceptance efficiency, $\epsilon^{\texttt{Sel}\cdot}$ is the selection efficiency and $\epsilon^{\texttt{PID}}$ is the PID selection efficiency. The geometrical acceptance efficiency accounts for the fraction of tracks that decay within the LHCb acceptance while the PID efficiency accounts for the PID selection requirements. The selection efficiency is evaluated by applying all of the remaining selection requirements simultaneously. It is possible to further divide the selection efficiency into the product of different sub-efficiencies which describe the different steps of the selection procedure. These additional sub-efficiencies include: the stripping and reconstruction efficiency which is the fraction of tracks which pass the stripping level requirements and are succesfully reconstructed into a b-hadron candidate; the trigger efficiency corresponding to the trigger lines outlined in Chapter 5; the efficiency of the q^2 requirement made on the dataset; the efficiency which accounts for the PID fiducial requirements and the track opening angle selection; the efficiency corresponding to the dipion mass, b-hadron mass and other invariant mass requirements made and, finally, the efficiency corresponding to the BDT requirement made. The total efficiency for a given mode can then be written in terms of these sub-efficiencies as

$$\begin{aligned}\epsilon = \epsilon^{\texttt{Geo.}} &\times \epsilon^{\texttt{Strip.}|\texttt{Geo.}} \times \epsilon^{\texttt{Trig.}|\texttt{Strip.\&Geo.}} \\ &\times \epsilon^{q^2|\texttt{Trig.\&Strip.\&Geo.}} \times \epsilon^{\texttt{Fid.}|q^2\&\texttt{Trig.\&Strip.\&Geo.}} \\ &\times \epsilon^{\texttt{Mass}|\texttt{Fid.}\&q^2\&\texttt{Trig.\&Strip.\&Geo}} \times \epsilon^{\texttt{BDT}|\texttt{Mass}\&\texttt{Fid.}\&q^2\&\texttt{Trig.\&Strip.\&Geo.}} \\ &\times \epsilon^{\texttt{PID}|\texttt{BDT}\&\texttt{Mass}\&\texttt{Fid.}\&q^2\&\texttt{Trig.}\&\texttt{Strip.}\&\texttt{Geo}} ,\end{aligned} \tag{7.3}$$

where $\epsilon^{\texttt{Strip}\cdot}$ is the stripping and reconstruction efficiency, $\epsilon^{\texttt{Trig}\cdot}$ is the trigger efficiency, ϵ^{q^2}

is the efficiency of the relevant q^2 selection, $\epsilon^{\texttt{Fid}\cdot}$ is the efficiency of the fiducial requirements, $\epsilon^{\texttt{Mass}}$ is the efficiency of the mass requirements and $\epsilon^{\texttt{BDT}}$ is the efficiency of the BDT requirement. As can be seen from Equation 7.3, the efficiencies are calculated sequentially so that, for example, the PID efficiency is calculated with the rest of the selection already applied. The various sub-efficiencies which make up the selection efficiency can be used to compare the effect of each selection step on each decay mode of interest in the analysis. However, when evaluating the final branching fractions, the efficiency definition used is that from Equation 7.1 where the selection sub-efficiencies are merged together and come with a single corresponding uncertainty. This is a preferable statistic to work with when performing final calculations because the single efficiency avoids correlations which may arise between uncertainties of the different sub-efficiencies. The geometrical and PID efficiencies are still kept separate as they are calculated from independent generator level and calibration samples.

Apart from the PID efficiency, the statistical uncertainties on the efficiencies defined within this chapter are calculated using a bootstrapping technique [191]. Using this approach, for each individual bootstrap and prior to the application of the requirement in question, each candidate is weighted with a randomly generated value taken from a Poisson distribution with a mean of 1. The effect of the application of the weight is that effectively, for each bootstrap, only a subset of the candidates are included in the calculation of the efficiency. When repeating the efficiency calculation for n bootstraps where n is large, the resulting efficiency distribution is expected to vary according to a Gaussian distribution. The mean of the Gaussian is taken as the efficiency value and the width is taken as the corresponding statistical uncertainty. An example of the bootstrapping procedure with the characteristic Gaussian distribution can be seen in Figure 7.1 which shows the results of the bootstrapping when calculating the trigger efficiency for the Run 1 $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC sample.

To determine the uncertainty of the PID selection efficiency a slightly different approach is taken. This alternative method involves the re-sampling of the LHCb calibration samples which are used to determine the PID efficiency. For each re-sampling, each bin of the calibration sample has its efficiency fluctuated within its uncertainty. Then, the total PID efficiency is recalculated by multiplying together the individual PID efficiencies of the four separate particle tracks. This exercise is repeated a large number of times and the resulting spread of PID efficiencies is seen to form a Gaussian distribution. The final PID efficiency is then taken as the mean of the distribution with the corresponding statistical uncertainty taken as the standard deviation.

In the following sections of this chapter more specific information is provided covering some of the different methods used to evaluate some of the sub-efficiencies. In addition, the corrections made to the MC concerning the trigger efficiencies, tracking efficiencies and stripping-level muon PID selections are introduced and explained. At the end of this chapter, the full set

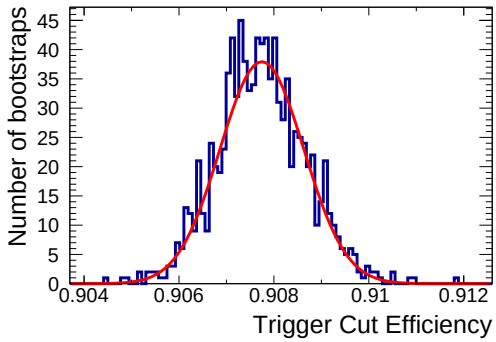

Figure 7.1: Example of the output of the bootstrapping procedure used to evaluate the trigger efficiency for the Run 1 $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC sample. The trigger efficiency distribution is fit with a Gaussian.

of calculated efficiencies are tabulated. All of the efficiency values are presented for the three datasets separately. For a given dataset the efficiency is calculated using the two combined years which make up the dataset meaning that, for example, the 2011 and 2012 candidates together are used to determine the average Run 1 efficiency. This is achieved by combining the two separate efficiencies via a weighted average using the integrated luminosity information.

7.2 Geometrical and Stripping Efficiencies

The geometrical acceptance efficiency is obtained using generator level MC samples. Following the generation of these separate samples, the efficiency is evaluated by applying a requirement on each final-state particle of $\theta \in [0.01, 0.4]$ mrad, where θ is defined as the angle between the particle three momentum and the beam axis direction. Most LHCb physics analyses evaluate the geometrical acceptance efficiencies using a set of collaboration-wide lookup tables which provide this information for each specific MC sample. However, by calculating this efficiency manually in this way it is possible to see how these numbers change for each region of q^2.

Similarly to the geometrical efficiency, the stripping and reconstruction efficiency is evaluated using a separate set of MC samples which have yet to pass through the stripping selection. The efficiency is determined by evaluating the fraction of events found in the pre-stripping MC samples which are also present within the post-stripping MC samples.

7.3 Trigger Efficiency Correction

As discussed in Chapter 3, the LHCb trigger system is composed of a hardware (L0) and software (HLT) stage. The ability of the LHCb MC samples to correctly model the efficiency attributed to the HLT trigger lines is good, partially due to the flexibility in modelling a purely software-based process. This is not necessarily the case however for the hardware-based L0 trigger where simply using the default MC samples to evaluate the L0 trigger efficiencies could result in the introduction of inaccuracies to the analysis. For example, the L0 trigger thresholds were changed over the course of a data-taking run but have a fixed value in the Run 1 simulation samples. As a result it is a standard practice within physics analyses to compare the L0 trigger efficiency between the data and MC and make corrections to the MC samples if the disagreement is large.

In this analysis, in order to correct the L0 trigger efficiency, the MC samples receive a set of correcting weights. As was done for the kinematic weighting described in Chapter 5 the weights are calculated by comparing the L0 trigger efficiency evaluated with MC simulation to that evaluated with the data. In order to extract the L0 trigger efficiencies from data a method is used which exploits events that have been triggered independently of the signal (TIS). The efficiency to trigger on the signal (TOS) is

$$\epsilon_{\text{TOS}} = \frac{N_{\text{TISTOS}}}{N_{\text{TIS}}} \,, \tag{7.4}$$

where N_{TIS} is the number of TIS candidates at L0 and N_{TISTOS} is the number of candidates that simultaneously satisfy the L0 `TOS` requirements. In this analysis, the L0 requirements of interest are the `L0Muon` and `L0Dimuon` trigger lines and the two corresponding efficiencies are compared between MC and data in bins of maximum muon p_T and the product of muon track p_T values, respectively.

To perform the study, candidates are taken from the $B^0 \to J/\psi K^{*0}$ control mode in the data and corresponding control mode MC samples. For the Run 1 study, for the HLT1 stage, the `Hlt1TrackAllL0Decision` line is required to be TOS. In addition, for the HLT 2 stage, either the `Hlt2Topo2BodyBBDTDecision`, `Hlt2Topo3BodyBBDTDecision` or `Hlt2Topo4BodyBBDTDecision` lines are required to be TOS. Only these HLT1 and HLT2 lines are used for the study because all other lines only consider events that have already passed `L0Muon`, meaning that their inclusion would bias the efficiency. The equivalent HLT1 and HLT2 lines used for the Run 2 study are `Hlt1TrackMVADecision` for HLT1 and either `Hlt2Topo2BodyDecision`, `Hlt2Topo3BodyDecision` or `Hlt2Topo4BodyDecision` for Run 2. Furthermore, either `L0HadronDecision` or `L0MuonDecision` are required to satisfy the TIS condition as most events which pass the L0 stage do so via one of these two

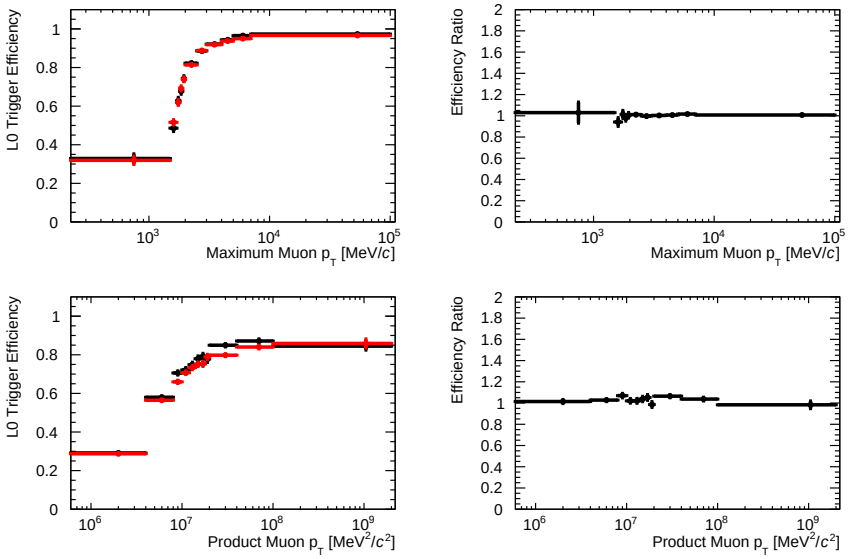

Figure 7.2: L0 trigger efficiency calculated for 2011 using MC and data via the TISTOS method. Left: The comparison between the data and MC L0 trigger efficiencies for the `L0Muon` (top) and `L0DiMuon` (bottom) lines. Right: The ratio between the data and MC efficiencies.

lines. As was mentioned in Chapter 5 events with nSPDHits$>$ 450 do not pass the `L0Muon` trigger in Run 2. Ideally, when testing the `L0DiMuon` efficiency, `L0DiMuon` should also be TIS. However, as too few candidates come through this line alone the resulting dataset is too small to proceed with the study. Therefore, it is assumed that the `L0DiMuon` TOS efficiency shows the same behaviour for nSPDHits$>$ 450 and nSPDHits$<$ 450.

The MC-data comparison is performed in 2D, testing both the `L0Muon` and `L0DiMuon` agreement simultaneously. However, to provide an initial proof of concept, the MC-data agreement is first checked for the `L0Muon` and `L0DiMuon` lines separately using a 1D binning scheme. An example of the 1D study, for 2011, is shown in Figure 7.2 which shows the L0 trigger efficiency for Run 1 for data and MC. As can be seen from this example the difference in data and MC efficiencies is small and the correcting weights in the bins are close in value to unity. The correcting weights are finally calculated using the full 2D binning scheme with the resultant histograms and weights, for each year of data-taking, shown in Figure 7.3.

7.4 Tracking Efficiency Correction

Another correction is made to the MC to account for the imperfect modelling of the track reconstruction efficiency in the simulation. These corrections are calculated centrally by the

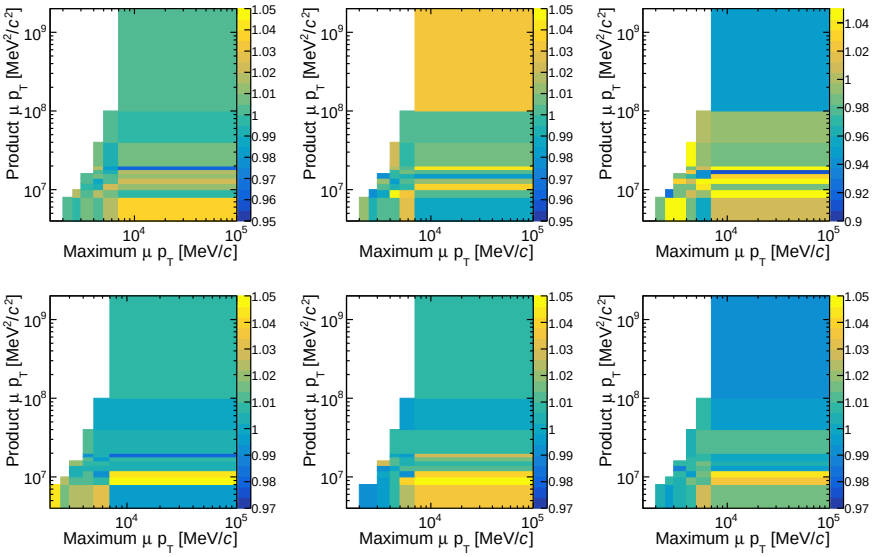

Figure 7.3: The L0 trigger efficiency ratio between data and MC binned in maximum muon p_{T} and the product of the muon p_{T}. The figures show, from the left to right on the top row the 2011, 2012 and 2015 histograms. On the bottom row the equivalents are shown for 2016, 2017 and 2018. The size of the correcting weights is seen to be approximately unity across the different histograms. The equivalent plots for the data and MC efficiencies can be found in Appendix C.

collaboration and are determined using a tag and probe technique with candidates originating from the decay of a J/ψ resonance to a dimuon final state. The technique fully reconstructs one track while partially reconstructing the other before determining how often the partially reconstructed track can then be fully reconstructed. Any tracking efficiency differences between hadronic and muonic tracks will be small and cancel within the final branching fraction ratio. Additionally, any differences in tracking efficiency between pion or kaon tracks which may not cancel within the ratio are known to be small [192].

The correcting weights are applied individually to each track in the samples. The correcting weights are derived from 2D histograms, which are binned in the track momentum and pseudorapidity, with the 2012, 2015 and 2016 tables shown in Figure 7.4. Due to the unavailability of the 2011, 2017 and 2018 tables at the time of the analysis the 2011 correction procedure uses the 2012 tables while the 2017 and 2018 corrections are made with the 2016 tables. This assumption is based on the fact that the tracking reconstruction procedure is unchanged between these years and so any differences are expected to be small. As can be seen from the tables the weights are all close to unity and, therefore, the resulting correction is very small.

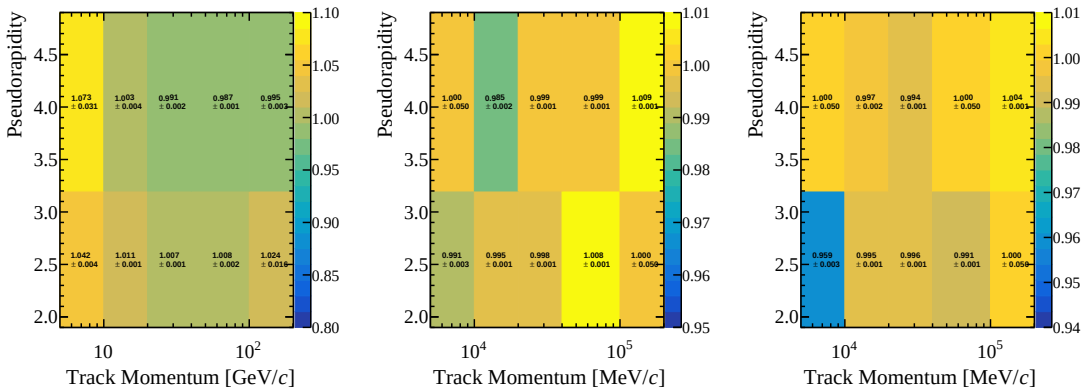

Figure 7.4: Tracking efficiency correction weights in bins of track momenta and pseudorapidity. Histograms are created centrally by the LHCb collaboration. The left most figure shows the 2012 histogram, the middle the 2015 histogram and the right the 2016 histogram.

7.5 Stripping PID Correction

An additional correction to the MC samples is also made to account for the loose muon PID requirements which exist at stripping level. These requirements are that the muon tracks must pass both the `IsMuon` and $\mathrm{DLL}_{\mu\pi} > -3$ conditions. The procedure is entirely analogous to the weighting of the MC samples using the `PIDCalib` calibration samples to determine the effect of the PID requirements made within the selection. The correcting weights are determined by comparing `PIDCalib` calibration samples containing muon tracks, which evaluate the efficiency of the muon PID requirements, with a separate, muon MC sample. The efficiencies of these two samples are compared using a 3D binning scheme which bins in the momentum, pseudorapidity and the number of particle tracks per event. Figure 7.5 shows an example of the calculated weights for Run 1, for a single slice in the occupancy variable which corresponds to nTracks $\in$ [0,100]. As can be seen the weights are close to unity meaning that the correction is very small.

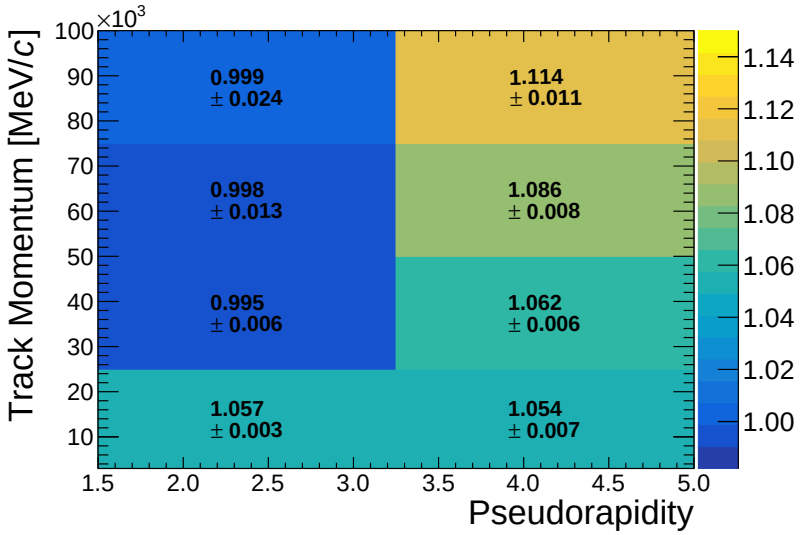

Figure 7.5: Ratio of stripping-level muon PID efficiencies between data and MC as a function of the muon momentum and pseudorapidity for $0 <$nTracks< 100. The weights are close to unity meaning that the correcting effect is small. Similar plots are seen for different slices in occupancy and for the Run 2 process.

7.6 Efficiency Tables

In this section the different efficiencies are presented for the control mode, resonant modes and rare modes. The efficiencies included within the tables include both the full range of different sub-efficiencies which are included for comparison and also the final efficiency values which are used to perform calculations throughout the analysis. Table 7.1 shows the sub-efficiency values for both the control mode and the resonant mode while Table 7.2 shows the corresponding final efficiency values. Both of these tables have some values which are presented twice; once for the case where a constraint is applied to set the dimuon mass to the mass of the J/ψ resonance, and once where this constraint is not applied. Both of these sets of numbers are used within Chapter 8 to perform the final resonant mode fits. No J/ψ mass constraint is included for the control mode as only a single fit is performed to this sample.

The efficiencies for the rare modes are presented under two different scenarios. Firstly, Table 7.3 and Table 7.4 show the sub-efficiencies and total efficiency values respectively. These values are all calculated using candidates which fall across the full $m(\pi^+\pi^-)$ and q^2 spectrum and therefore include the efficiency of the $m(\pi^+\pi^-)$ and q^2 windows. This means, for example, that the effect of the $m(\pi^+\pi^-)$ requirement applied during the selection, only affects those sub-efficiencies from the mass efficiency onwards in the definition from Equation 7.3 (i.e. the mass, BDT and PID efficiencies.). These efficiency values are not used to calculate the final rare mode branching fractions, but are instead used throughout the analysis to estimate the yield of different modes under different sets of requirements, such as for the pseudo-experiment study referred to in Chapter 8. As can be seen from the tables, these values are calculated for the three different q^2 regimes of interest, full q^2, low q^2 and high q^2.

The second set of rare mode efficiencies can be found in Table 7.5 and Table 7.6 which contain the sub-efficiency break down and the total efficiencies respectively. These values are calculated using candidates which fall into the low $m(\pi^+\pi^-)$ bin ($m(\pi^+\pi^-) \in [515, 1300]\,\mathrm{MeV}/c^2$) and relevant q^2 bin exclusively. This means that all of the efficiencies are calculated with the relevant $m(\pi^+\pi^-)$ and q^2 requirement already applied. This also means that the values of the q^2 efficiency requirement, for both the low and high q^2 regions is, by definition, unity within these tables. The full q^2 efficiency result uses the q^2 sub-efficiency to extrapolate the branching fraction result to the full q^2 range. It is this set of efficiency values which are used to provide the final branching fraction measurements found in Chapter 10.

Decay	Dataset	$\epsilon^{\texttt{Strip.}}$	$\epsilon^{\texttt{Trig.}}$	ϵ^{q^2}	$\epsilon^{\texttt{Fid.}}$	$\epsilon^{\texttt{Mass}}$		$\epsilon^{\texttt{BDT}}$	
						No J/ψ constraint	J/ψ constraint	No J/ψ constraint	J/ψ constraint
$B^0 \to J/\psi\pi^+\pi^-$	2011+2012	0.0988 ± 0.0005	0.9056 ± 0.0016	0.9962 ± 0.0002	0.7453 ± 0.0025	0.7868 ± 0.0038	0.8054 ± 0.0038	0.5917 ± 0.0022	0.5899 ± 0.0022
	2015+2016	0.1079 ± 0.0010	0.7324 ± 0.0039	0.9953 ± 0.0003	0.7017 ± 0.0056	0.7622 ± 0.0078	0.7793 ± 0.0080	0.5749 ± 0.0042	0.5724 ± 0.0043
	2017+2018	0.1122 ± 0.0010	0.8053 ± 0.0030	0.9950 ± 0.0006	0.7060 ± 0.0044	0.7710 ± 0.0098	0.7876 ± 0.0102	0.5868 ± 0.0034	0.5845 ± 0.0034
$B_s^0 \to J/\psi\pi^+\pi^-$	2011+2012	0.1125 ± 0.0006	0.9167 ± 0.0015	0.9941 ± 0.0004	0.7609 ± 0.0025	0.7444 ± 0.0044	0.7505 ± 0.0043	0.6001 ± 0.0023	0.5990 ± 0.0023
	2015+2016	0.1256 ± 0.0011	0.7293 ± 0.0041	0.9933 ± 0.0005	0.7058 ± 0.0046	0.7104 ± 0.0076	0.7170 ± 0.0073	0.5881 ± 0.0038	0.5865 ± 0.0037
	2017+2018	0.1281 ± 0.0010	0.7996 ± 0.0039	0.9932 ± 0.0006	0.7179 ± 0.0035	0.7411 ± 0.0058	0.7462 ± 0.0061	0.5908 ± 0.0043	0.5893 ± 0.0043
$B_s^0 \to J/\psi\eta'$	2011+2012	0.0529 ± 0.0002	0.9077 ± 0.0014	0.9981 ± 0.0002	0.7404 ± 0.0021	0.4577 ± 0.0027	0.5578 ± 0.0026	0.5494 ± 0.0041	0.5310 ± 0.0038
	2015+2016	0.0664 ± 0.0003	0.7348 ± 0.0019	0.9645 ± 0.0008	0.7016 ± 0.0021	0.4448 ± 0.0029	0.4592 ± 0.0030	0.5445 ± 0.0042	0.5429 ± 0.0042
	2017+2018	0.0690 ± 0.0001	0.8029 ± 0.0007	0.9644 ± 0.0003	0.7035 ± 0.0008	0.4520 ± 0.0010	0.4675 ± 0.0011	0.5502 ± 0.0015	0.5484 ± 0.0015
$B_s^0 \to J/\psi\phi$	2011+2012	0.0326 ± 0.0001	0.8976 ± 0.0012	0.9722 ± 0.0006	0.7477 ± 0.0018	0.1465 ± 0.0017	0.2414 ± 0.0020	0.5700 ± 0.0061	0.5434 ± 0.0052
	2015+2016	0.0526 ± 0.0003	0.7773 ± 0.0023	0.9604 ± 0.0011	0.7099 ± 0.0021	0.1056 ± 0.0018	0.1749 ± 0.0024	0.5533 ± 0.0095	0.5085 ± 0.0075
	2017+2018	0.0384 ± 0.0003	0.8025 ± 0.0034	0.9671 ± 0.0014	0.7024 ± 0.0033	0.1455 ± 0.0031	0.1522 ± 0.0033	0.5563 ± 0.0160	0.5481 ± 0.0164
$B^0 \to J/\psi(K^+ \to \pi^+)\pi^-$	2011+2012	0.1132 ± 0.0006	0.9004 ± 0.0017	0.9965 ± 0.0003	0.8008 ± 0.0023	0.3274 ± 0.0029	0.4218 ± 0.0030	0.5941 ± 0.0051	0.5942 ± 0.0044
	2015+2016	0.1189 ± 0.0008	0.7811 ± 0.0056	0.9944 ± 0.0016	0.7592 ± 0.0047	0.3267 ± 0.0065	0.3405 ± 0.0065	0.5846 ± 0.0121	0.5881 ± 0.0119
	2017+2018	0.1277 ± 0.0019	0.8307 ± 0.0063	0.9957 ± 0.0007	0.7457 ± 0.0072	0.3367 ± 0.0084	0.4358 ± 0.0081	0.5654 ± 0.0162	0.5723 ± 0.0143
$B^0 \to J/\psi K^{*0}$	2011+2012	0.1010 ± 0.0001	0.9033 ± 0.0004	0.9954 ± 0.0001	0.7735 ± 0.0006	0.8001 ± 0.0006	-	0.6005 ± 0.0009	-
	2015+2016	0.1058 ± 0.0003	0.7317 ± 0.0016	0.9950 ± 0.0002	0.7358 ± 0.0017	0.8017 ± 0.0017	-	0.5878 ± 0.0023	-
	2017+2018	0.1150 ± 0.0002	0.7983 ± 0.0008	0.9954 ± 0.0001	0.7319 ± 0.0014	0.7996 ± 0.0008	-	0.6023 ± 0.0010	-

Table 7.1: Calculated sub-efficiencies for the control mode and resonant modes where the quoted uncertainty is statistical. Values are presented for the cases where a J/ψ mass constraint is present and also where it is absent.

Decay	Dataset	$\epsilon^{\texttt{Geo.}}$	$\epsilon^{\texttt{Selection}}$		$\epsilon^{\texttt{PID}}$		$\epsilon^{\texttt{Tot.}}$	
			No J/ψ constraint	J/ψ constraint	No J/ψ constraint	J/ψ constraint	No J/ψ constraint	J/ψ constraint
$B^0 \to J/\psi\,\pi^+\pi^-$	2011+2012	0.1433 ± 0.0004	0.0309 ± 0.0002	0.0315 ± 0.0002	0.7589 ± 0.0034	0.7588 ± 0.0034	$(3.36 \pm 0.22) \times 10^{-3}$	$(3.43 \pm 0.03) \times 10^{-3}$
	2015+2016	0.1495 ± 0.0005	0.0249 ± 0.0002	0.0253 ± 0.0002	0.7697 ± 0.0020	0.7698 ± 0.0020	$(2.87 \pm 0.03) \times 10^{-3}$	$(2.91 \pm 0.03) \times 10^{-3}$
	2017+2018	0.1508 ± 0.0005	0.0292 ± 0.0003	0.0297 ± 0.0003	0.8136 ± 0.0082	0.8134 ± 0.0079	$(3.58 \pm 0.05) \times 10^{-3}$	$(3.64 \pm 0.05) \times 10^{-3}$
$B_s^0 \to J/\psi\,\pi^+\pi^-$	2011+2012	0.1409 ± 0.0004	0.0349 ± 0.0002	0.0351 ± 0.0002	0.7585 ± 0.0035	0.7585 ± 0.0035	$(3.73 \pm 0.03) \times 10^{-3}$	$(3.75 \pm 0.03) \times 10^{-3}$
	2015+2016	0.1469 ± 0.0005	0.0277 ± 0.0003	0.0279 ± 0.0003	0.7726 ± 0.0020	0.7727 ± 0.0020	$(3.14 \pm 0.04) \times 10^{-3}$	$(3.17 \pm 0.03) \times 10^{-3}$
	2017+2018	0.1486 ± 0.0005	0.0326 ± 0.0003	0.0327 ± 0.0003	0.8136 ± 0.0064	0.8138 ± 0.0063	$(3.94 \pm 0.05) \times 10^{-3}$	$(3.95 \pm 0.05) \times 10^{-3}$
$B_s^0 \to J/\psi\,\eta'$	2011+2012	0.1445 ± 0.0005	0.0090 ± 0.0001	0.0105 ± 0.0001	0.7696 ± 0.0026	0.7719 ± 0.0026	$(1.00 \pm 0.02) \times 10^{-3}$	$(1.17 \pm 0.05) \times 10^{-3}$
	2015+2016	0.1492 ± 0.0005	0.0082 ± 0.0001	0.0085 ± 0.0001	0.7798 ± 0.0021	0.7799 ± 0.0020	$(0.95 \pm 0.01) \times 10^{-3}$	$(0.99 \pm 0.01) \times 10^{-3}$
	2017+2018	0.1514 ± 0.0005	0.0095 ± 0.0001	0.0098 ± 0.0001	0.8208 ± 0.0067	0.8209 ± 0.0068	$(1.18 \pm 0.02) \times 10^{-3}$	$(1.22 \pm 0.02) \times 10^{-3}$
$B_s^0 \to J/\psi\,\phi$	2011+2012	0.1480 ± 0.0004	0.0018 ± 0.0001	0.0028 ± 0.0001	0.7630 ± 0.0032	0.7665 ± 0.0031	$(0.20 \pm 0.01) \times 10^{-3}$	$(0.32 \pm 0.01) \times 10^{-3}$
	2015+2016	0.1542 ± 0.0005	0.0016 ± 0.0001	0.0025 ± 0.0001	0.7978 ± 0.0009	0.8022 ± 0.0009	$(0.20 \pm 0.01) \times 10^{-3}$	$(0.31 \pm 0.01) \times 10^{-3}$
	2017+2018	0.1557 ± 0.0005	0.0017 ± 0.0001	0.0018 ± 0.0001	0.8217 ± 0.0248	0.8145 ± 0.0243	$(0.22 \pm 0.01) \times 10^{-3}$	$(0.23 \pm 0.01) \times 10^{-3}$
$B^0 \to J/\psi\,(K^+ \to \pi^+)\pi^-$	2011+2012	0.1478 ± 0.0004	0.0158 ± 0.0002	0.02024 ± 0.0002	0.0153 ± 0.0004	0.0143 ± 0.0006	$(0.36 \pm 0.01) \times 10^{-4}$	$(0.43 \pm 0.02) \times 10^{-4}$
	2015+2016	0.1545 ± 0.0005	0.0178 ± 0.0012	0.0213 ± 0.0011	0.0084 ± 0.0001	0.0082 ± 0.0001	$(0.23 \pm 0.02) \times 10^{-4}$	$(0.27 \pm 0.01) \times 10^{-4}$
	2017+2018	0.1552 ± 0.0005	0.0158 ± 0.0005	0.0208 ± 0.0008	0.0113 ± 0.0001	0.0118 ± 0.0001	$(0.28 \pm 0.03) \times 10^{-4}$	$(0.38 \pm 0.15) \times 10^{-4}$
$B^0 \to J/\psi\,K^{*0}$	2011+2012	0.1489 ± 0.0004	0.0337 ± 0.0001	-	0.7726 ± 0.0037	-	$(3.88 \pm 0.02) \times 10^{-3}$	-
	2015+2016	0.1555 ± 0.0005	0.0276 ± 0.0001	-	0.6923 ± 0.0019	-	$(2.97 \pm 0.02) \times 10^{-3}$	-
	2017+2018	0.1565 ± 0.0004	0.0329 ± 0.0001	-	0.7032 ± 0.0060	-	$(3.62 \pm 0.03) \times 10^{-3}$	-

Table 7.2: Calculated final efficiency values for the control mode and resonant modes where the quoted uncertainty is statistical. Values are presented for the cases where a J/ψ mass constraint is present and also where it is absent.

Decay	q^2 Range	Dataset	$\epsilon^{\texttt{Strip.}}$	$\epsilon^{\texttt{Trig.}}$	ϵ^{q^2}	$\epsilon^{\texttt{Fid.}}$	$\epsilon^{\texttt{Mass}}$	$\epsilon^{\texttt{BDT}}$
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.0975 ± 0.0003	0.9077 ± 0.0009	0.6768 ± 0.0014	0.7214 ± 0.0017	0.7258 ± 0.0027	0.5920 ± 0.0025
	Full	2015+2016	0.1076 ± 0.0008	0.6979 ± 0.0037	0.6552 ± 0.0041	0.6690 ± 0.0054	0.7325 ± 0.0072	0.5917 ± 0.0090
	Full	2017+2018	0.1127 ± 0.0005	0.7657 ± 0.0028	0.6647 ± 0.0026	0.6788 ± 0.0029	0.7188 ± 0.0047	0.5968 ± 0.0040
	Low	2011+2012	0.0975 ± 0.0003	0.9077 ± 0.0009	0.4368 ± 0.0017	0.7590 ± 0.0022	0.6574 ± 0.0037	0.5940 ± 0.0031
	Low	2015+2016	0.1076 ± 0.0008	0.6979 ± 0.0037	0.3971 ± 0.0049	0.7114 ± 0.0064	0.6534 ± 0.0101	0.5992 ± 0.0134
	Low	2017+2018	0.1127 ± 0.0005	0.7657 ± 0.0028	0.4238 ± 0.0031	0.7209 ± 0.0036	0.6447 ± 0.0066	0.5957 ± 0.0052
	High	2011+2012	0.0975 ± 0.0003	0.9077 ± 0.0009	0.1499 ± 0.0010	0.6033 ± 0.0034	0.9263 ± 0.0031	0.5827 ± 0.0045
	High	2015+2016	0.1076 ± 0.0008	0.6979 ± 0.0037	0.1670 ± 0.0033	0.5655 ± 0.0087	0.9216 ± 0.0075	0.5758 ± 0.0125
	High	2017+2018	0.1127 ± 0.0005	0.7657 ± 0.0028	0.1525 ± 0.0016	0.5558 ± 0.0058	0.9314 ± 0.0042	0.6030 ± 0.0075
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.1065 ± 0.0004	0.9046 ± 0.0011	0.7228 ± 0.0015	0.7707 ± 0.0017	0.6716 ± 0.0035	0.5983 ± 0.0014
	Full	2015+2016	0.1221 ± 0.0011	0.6686 ± 0.0044	0.6995 ± 0.0044	0.7204 ± 0.0052	0.6789 ± 0.0121	0.6042 ± 0.0039
	Full	2017+2018	0.1287 ± 0.0008	0.7450 ± 0.0036	0.7237 ± 0.0027	0.7215 ± 0.0034	0.6699 ± 0.0064	0.6037 ± 0.0024
	Low	2011+2012	0.1065 ± 0.0004	0.9046 ± 0.0011	0.6059 ± 0.0019	0.7831 ± 0.0020	0.6437 ± 0.0034	0.6007 ± 0.0014
	Low	2015+2016	0.1221 ± 0.0011	0.6686 ± 0.0044	0.5679 ± 0.0056	0.7305 ± 0.0056	0.6561 ± 0.0110	0.6072 ± 0.0047
	Low	2017+2018	0.1287 ± 0.0008	0.7450 ± 0.0036	0.6056 ± 0.0031	0.7333 ± 0.0041	0.6389 ± 0.0070	0.6026 ± 0.0026
	High	2011+2012	0.1065 ± 0.0004	0.9046 ± 0.0011	0.0423 ± 0.0004	0.6442 ± 0.0040	0.9759 ± 0.0021	0.5931 ± 0.0046
	High	2015+2016	0.1221 ± 0.0011	0.6686 ± 0.0044	0.0476 ± 0.0011	0.6069 ± 0.0096	0.9654 ± 0.0052	0.5762 ± 0.0137
	High	2017+2018	0.1287 ± 0.0008	0.7450 ± 0.0036	0.0446 ± 0.0007	0.6013 ± 0.0062	0.9698 ± 0.0046	0.6108 ± 0.0092
$B^0 \to (K^+ \to \pi^+)\pi^-\mu^+\mu^-$	Full	2011+2012	0.0972 ± 0.0016	0.8918 ± 0.0051	0.6964 ± 0.0110	0.7722 ± 0.0053	0.3279 ± 0.0074	0.5785 ± 0.0112
	Full	2015+2016	0.1037 ± 0.0013	0.6911 ± 0.0061	0.6871 ± 0.0079	0.7235 ± 0.0088	0.2948 ± 0.0079	0.6111 ± 0.0103
	Full	2017+2018	0.1080 ± 0.0010	0.7506 ± 0.0041	0.7036 ± 0.0031	0.7212 ± 0.0058	0.3279 ± 0.0052	0.6068 ± 0.0062
	Low	2011+2012	0.0972 ± 0.0016	0.8918 ± 0.0051	0.5182 ± 0.0098	0.8038 ± 0.0065	0.3489 ± 0.0100	0.5778 ± 0.0135
	Low	2015+2016	0.1037 ± 0.0013	0.6911 ± 0.0061	0.4999 ± 0.0085	0.7582 ± 0.0115	0.3096 ± 0.0110	0.6183 ± 0.0114
	Low	2017+2018	0.1080 ± 0.0010	0.7506 ± 0.0041	0.5179 ± 0.0048	0.7528 ± 0.0077	0.3495 ± 0.0070	0.6050 ± 0.0068
	High	2011+2012	0.0972 ± 0.0016	0.8918 ± 0.0051	0.1006 ± 0.0021	0.6262 ± 0.0058	0.1935 ± 0.0064	0.5824 ± 0.0204
	High	2015+2016	0.1037 ± 0.0013	0.6911 ± 0.0061	0.1067 ± 0.0024	0.5769 ± 0.0076	0.1915 ± 0.0084	0.5998 ± 0.0280
	High	2017+2018	0.1080 ± 0.0010	0.7506 ± 0.0041	0.1018 ± 0.0014	0.5683 ± 0.0048	0.1870 ± 0.0036	0.6214 ± 0.0100

Table 7.3: Calculated sub-efficiencies for the rare mode processes for the different q^2 regimes. The quoted uncertainty is statistical. All values presented are calculated using candidates taken from the full q^2 and $m(\pi^+\pi^-)$ range.

Decay	q^2 Range	Dataset	$\epsilon^{\texttt{Geo.}}$	$\epsilon^{\texttt{Sel.}}$	$\epsilon^{\texttt{PID}}$	$\epsilon^{\texttt{Tot.}}$
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.1459 ± 0.0004	0.0186 ± 0.0002	0.7524 ± 0.0034	$(2.04 \pm 0.02) \times 10^{-3}$
	Full	2015+2016	0.1518 ± 0.0005	0.0145 ± 0.0003	0.7731 ± 0.0022	$(1.70 \pm 0.03) \times 10^{-3}$
	Full	2017+2018	0.1531 ± 0.0005	0.0169 ± 0.0003	0.8095 ± 0.0083	$(2.10 \pm 0.03) \times 10^{-3}$
	Low	2011+2012	0.1459 ± 0.0004	0.0119 ± 0.0001	0.7249 ± 0.0033	$(1.26 \pm 0.01) \times 10^{-3}$
	Low	2015+2016	0.1518 ± 0.0005	0.0085 ± 0.0002	0.7493 ± 0.0021	$(9.63 \pm 0.22) \times 10^{-4}$
	Low	2017+2018	0.1531 ± 0.0005	0.0103 ± 0.0003	0.7919 ± 0.0137	$(1.25 \pm 0.03) \times 10^{-3}$
	High	2011+2012	0.1459 ± 0.0004	0.0043 ± 0.0001	0.7932 ± 0.0043	$(5.01 \pm 0.06) \times 10^{-4}$
	High	2015+2016	0.1518 ± 0.0005	0.0038 ± 0.0001	0.8077 ± 0.0026	$(4.66 \pm 0.15 \times 10^{-4}$
	High	2017+2018	0.1531 ± 0.0005	0.0041 ± 0.0001	0.8367 ± 0.0018	$(5.28 \pm 0.11 \times 10^{-4}$
$B^0_s \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.1421 ± 0.0004	0.0215 ± 0.0002	0.7401 ± 0.0034	$(2.26 \pm 0.01) \times 10^{-3}$
	Full	2015+2016	0.1486 ± 0.0005	0.0172 ± 0.0002	0.7611 ± 0.0021	$(1.94 \pm 0.02) \times 10^{-3}$
	Full	2017+2018	0.1501 ± 0.0005	0.0206 ± 0.0004	0.8027 ± 0.0109	$(2.48 \pm 0.04) \times 10^{-3}$
	Low	2011+2012	0.1421 ± 0.0004	0.0176 ± 0.0001	0.7309 ± 0.0034	$(1.83 \pm 0.01) \times 10^{-3}$
	Low	2015+2016	0.1486 ± 0.0005	0.0138 ± 0.0002	0.7533 ± 0.0021	$(1.54 \pm 0.02) \times 10^{-3}$
	Low	2017+2018	0.1501 ± 0.0005	0.0166 ± 0.0004	0.7970 ± 0.0129	$(1.99 \pm 0.04) \times 10^{-3}$
	High	2011+2012	0.1421 ± 0.0004	0.0015 ± 0.0001	0.7907 ± 0.0041	$(1.70 \pm 0.02) \times 10^{-4}$
	High	2015+2016	0.1486 ± 0.0005	0.0013 ± 0.0001	0.7986 ± 0.0027	$(1.58 \pm 0.05) \times 10^{-4}$
	High	2017+2018	0.1501 ± 0.0005	0.0015 ± 0.0001	0.8355 ± 0.0023	$(1.92 \pm 0.05) \times 10^{-4}$
$B^0 \to (K^+ \to \pi^+)\pi^-\mu^+\mu^-$	Full	2011+2012	0.1518 ± 0.0004	0.0091 ± 0.0003	0.0165 ± 0.0004	$(2.27 \pm 0.05) \times 10^{-5}$
	Full	2015+2016	0.1582 ± 0.0005	0.0067 ± 0.0002	0.0078 ± 0.0001	$(8.24 \pm 0.21) \times 10^{-6}$
	Full	2017+2018	0.1598 ± 0.0006	0.0084 ± 0.0003	0.0119 ± 0.0003	$(1.60 \pm 0.03) \times 10^{-5}$
	Low	2011+2012	0.1518 ± 0.0004	0.0072 ± 0.0002	0.0180 ± 0.0003	$(1.97 \pm 0.05) \times 10^{-5}$
	Low	2015+2016	0.1582 ± 0.0005	0.0052 ± 0.0002	0.0081 ± 0.0001	$(6.66 \pm 0.19) \times 10^{-6}$
	Low	2017+2018	0.1598 ± 0.0006	0.0068 ± 0.0003	0.0130 ± 0.0004	$(1.42 \pm 0.03) \times 10^{-5}$
	High	2011+2012	0.1518 ± 0.0004	0.0006 ± 0.0001	0.0093 ± 0.0006	$(8.40 \pm 1.00) \times 10^{-7}$
	High	2015+2016	0.1582 ± 0.0005	0.0005 ± 0.0001	0.0059 ± 0.0002	$(4.84 \pm 0.20) \times 10^{-7}$
	High	2017+2018	0.1598 ± 0.0006	0.0005 ± 0.0001	0.0072 ± 0.0002	$(6.18 \pm 0.16) \times 10^{-7}$

Table 7.4: Calculated total efficiencies for the rare mode processes for the different q^2 regimes. The quoted uncertainty is statistical. All values presented are calculated using candidates taken from the full q^2 and $m(\pi^+\pi^-)$ range. The presented values are used to calculate expected yields within the analysis.

Decay	q^2 Range	Dataset	$\epsilon^{\texttt{Strip.}}$	$\epsilon^{\texttt{Trig.}}$	ϵ^{q^2}	$\epsilon^{\texttt{Fid.}}$	$\epsilon^{\texttt{Mass}}$	$\epsilon^{\texttt{BDT}}$
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.0960 ± 0.0003	0.9115 ± 0.0009	0.6634 ± 0.0014	0.7010 ± 0.0017	0.9517 ± 0.0011	0.5920 ± 0.0025
	Full	2015+2016	0.1050 ± 0.0008	0.7134 ± 0.0039	0.6470 ± 0.0041	0.6520 ± 0.0051	0.9472 ± 0.0028	0.5917 ± 0.0090
	Full	2017+2018	0.1101 ± 0.0005	0.7750 ± 0.0027	0.6509 ± 0.0026	0.6535 ± 0.0032	0.9491 ± 0.0019	0.5968 ± 0.0040
	Low	2011+2012	0.0916 ± 0.0004	0.8936 ± 0.0017	1.0000 ± 0.0000	0.7402 ± 0.0020	0.9438 ± 0.0016	0.5940 ± 0.0031
	Low	2015+2016	0.1042 ± 0.0014	0.6156 ± 0.0073	1.0000 ± 0.0000	0.6949 ± 0.0072	0.9361 ± 0.0041	0.5992 ± 0.0134
	Low	2017+2018	0.1092 ± 0.0009	0.7048 ± 0.0045	1.0000 ± 0.0000	0.6952 ± 0.0043	0.9417 ± 0.0028	0.5957 ± 0.0052
	High	2011+2012	0.0970 ± 0.0006	0.9337 ± 0.0017	1.0000 ± 0.0000	0.6010 ± 0.0033	0.9757 ± 0.0013	0.5827 ± 0.0045
	High	2015+2016	0.1018 ± 0.0017	0.8303 ± 0.0068	1.0000 ± 0.0000	0.5635 ± 0.0088	0.9748 ± 0.0029	0.5758 ± 0.0125
	High	2017+2018	0.1052 ± 0.0012	0.8515 ± 0.0057	1.0000 ± 0.0000	0.5543 ± 0.0057	0.9719 ± 0.0024	0.6030 ± 0.0075
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.1034 ± 0.0002	0.9079 ± 0.0005	0.7035 ± 0.0010	0.7593 ± 0.0011	0.9618 ± 0.0005	0.5983 ± 0.0014
	Full	2015+2016	0.1184 ± 0.0005	0.6807 ± 0.0024	0.6826 ± 0.0024	0.7069 ± 0.0028	0.9583 ± 0.0016	0.6042 ± 0.0039
	Full	2017+2018	0.1235 ± 0.0004	0.7569 ± 0.0017	0.6998 ± 0.0015	0.7113 ± 0.0018	0.9599 ± 0.0009	0.6037 ± 0.0024
	Low	2011+2012	0.0987 ± 0.0002	0.8984 ± 0.0007	1.0000 ± 0.0000	0.7750 ± 0.0012	0.9580 ± 0.0006	0.6007 ± 0.0014
	Low	2015+2016	0.1172 ± 0.0007	0.6221 ± 0.0030	1.0000 ± 0.0000	0.7204 ± 0.0033	0.9538 ± 0.0018	0.6072 ± 0.0047
	Low	2017+2018	0.1221 ± 0.0005	0.7147 ± 0.0022	1.0000 ± 0.0000	0.7263 ± 0.0020	0.9557 ± 0.0011	0.6026 ± 0.0026
	High	2011+2012	0.1070 ± 0.0007	0.9316 ± 0.0018	1.0000 ± 0.0000	0.6449 ± 0.0037	0.9937 ± 0.0008	0.5931 ± 0.0046
	High	2015+2016	0.1140 ± 0.0022	0.8279 ± 0.0075	1.0000 ± 0.0000	0.6065 ± 0.0091	0.9874 ± 0.0044	0.5762 ± 0.0137
	High	2017+2018	0.1208 ± 0.0014	0.8628 ± 0.0055	1.0000 ± 0.0000	0.6024 ± 0.0060	0.9919 ± 0.0017	0.6108 ± 0.0092
$B^0 \to (K^+ \to \pi^+)\pi^-\mu^+\mu^-$	Full	2011+2012	0.0937 ± 0.0010	0.9002 ± 0.0030	0.6879 ± 0.0051	0.7461 ± 0.0048	0.4995 ± 0.0083	0.5785 ± 0.0112
	Full	2015+2016	0.1013 ± 0.0012	0.6886 ± 0.0066	0.6825 ± 0.0061	0.7180 ± 0.0074	0.4305 ± 0.0137	0.6111 ± 0.0103
	Full	2017+2018	0.1048 ± 0.0007	0.7554 ± 0.0040	0.6876 ± 0.0029	0.6982 ± 0.0048	0.4985 ± 0.0047	0.6068 ± 0.0062
	Low	2011+2012	0.0988 ± 0.0019	0.8855 ± 0.0050	1.0000 ± 0.0000	0.7829 ± 0.0059	0.5533 ± 0.0101	0.5778 ± 0.0135
	Low	2015+2016	0.1144 ± 0.0023	0.6157 ± 0.0101	1.0000 ± 0.0000	0.7609 ± 0.0099	0.4673 ± 0.0182	0.6183 ± 0.0114
	Low	2017+2018	0.1195 ± 0.0013	0.6993 ± 0.0049	1.0000 ± 0.0000	0.7385 ± 0.0057	0.5454 ± 0.0069	0.6050 ± 0.0068
	High	2011+2012	0.0836 ± 0.0009	0.9318 ± 0.0025	1.0000 ± 0.0000	0.6071 ± 0.0052	0.2513 ± 0.0078	0.5824 ± 0.0204
	High	2015+2016	0.0808 ± 0.0010	0.8233 ± 0.0057	1.0000 ± 0.0000	0.5582 ± 0.0063	0.2506 ± 0.0106	0.5998 ± 0.0280
	High	2017+2018	0.0811 ± 0.0006	0.8630 ± 0.0032	1.0000 ± 0.0000	0.5418 ± 0.0045	0.2482 ± 0.0040	0.6214 ± 0.0100

Table 7.5: Calculated sub-efficiencies for the rare mode processes for the different q^2 regimes. The quoted uncertainty is statistical. All values presented are calculated using candidates taken from the low $m(\pi^+\pi^-)$ bin and from the relevant q^2 bin.

Decay	q^2 Range	Dataset	$\epsilon^{\texttt{Geo.}}$	$\epsilon^{\texttt{Sel.}}$	$\epsilon^{\texttt{PID}}$	$\epsilon^{\texttt{Tot.}}$
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.1459 ± 0.0004	0.0229 ± 0.0002	0.7524 ± 0.0034	$(2.51 \pm 0.02) \times 10^{-3}$
	Full	2015+2016	0.1520 ± 0.0005	0.0178 ± 0.0004	0.7731 ± 0.0022	$(2.09 \pm 0.04) \times 10^{-3}$
	Full	2017+2018	0.1534 ± 0.0005	0.0209 ± 0.0004	0.8095 ± 0.0083	$(2.59 \pm 0.04) \times 10^{-3}$
	Low	2011+2012	0.1456 ± 0.0006	0.0336 ± 0.0004	0.7249 ± 0.0033	$(3.55 \pm 0.04) \times 10^{-3}$
	Low	2015+2016	0.1535 ± 0.0008	0.0248 ± 0.0005	0.7493 ± 0.0021	$(2.85 \pm 0.06) \times 10^{-3}$
	Low	2017+2018	0.1535 ± 0.0009	0.0304 ± 0.0009	0.7919 ± 0.0137	$(3.69 \pm 0.09) \times 10^{-3}$
	High	2011+2012	0.1488 ± 0.0010	0.0302 ± 0.0005	0.7932 ± 0.0043	$(3.56 \pm 0.05) \times 10^{-3}$
	High	2015+2016	0.1525 ± 0.0012	0.0266 ± 0.0009	0.8077 ± 0.0026	$(3.28 \pm 0.11) \times 10^{-3}$
	High	2017+2018	0.1536 ± 0.0012	0.0286 ± 0.0020	0.8367 ± 0.0018	$(3.68 \pm 0.25) \times 10^{-3}$
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	Full	2011+2012	0.1423 ± 0.0004	0.0287 ± 0.0002	0.7401 ± 0.0034	$(3.02 \pm 0.02) \times 10^{-3}$
	Full	2015+2016	0.1487 ± 0.0005	0.0228 ± 0.0003	0.7611 ± 0.0021	$(2.58 \pm 0.03) \times 10^{-3}$
	Full	2017+2018	0.1501 ± 0.0005	0.0274 ± 0.0006	0.8027 ± 0.0109	$(3.30 \pm 0.05) \times 10^{-3}$
	Low	2011+2012	0.1427 ± 0.0005	0.0393 ± 0.0004	0.7309 ± 0.0034	$(4.10 \pm 0.04) \times 10^{-3}$
	Low	2015+2016	0.1481 ± 0.0007	0.0302 ± 0.0004	0.7533 ± 0.0021	$(3.37 \pm 0.04) \times 10^{-3}$
	Low	2017+2018	0.1505 ± 0.0006	0.0370 ± 0.0009	0.7970 ± 0.0129	$(4.44 \pm 0.08) \times 10^{-3}$
	High	2011+2012	0.1441 ± 0.0016	0.0369 ± 0.0007	0.7907 ± 0.0041	$(4.21 \pm 0.06) \times 10^{-3}$
	High	2015+2016	0.1550 ± 0.0022	0.0326 ± 0.0014	0.7986 ± 0.0027	$(4.04 \pm 0.16) \times 10^{-3}$
	High	2017+2018	0.1550 ± 0.0020	0.0375 ± 0.0011	0.8355 ± 0.0023	$(4.85 \pm 0.12) \times 10^{-3}$
$B^0 \to (K^+ \to \pi^+)\pi^-\mu^+\mu^-$	Full	2011+2012	0.1505 ± 0.0004	0.0127 ± 0.0004	0.0165 ± 0.0004	$(3.16 \pm 0.08) \times 10^{-5}$
	Full	2015+2016	0.1569 ± 0.0005	0.0094 ± 0.0003	0.0078 ± 0.0001	$(1.15 \pm 0.03) \times 10^{-5}$
	Full	2017+2018	0.1589 ± 0.0007	0.0118 ± 0.0004	0.0119 ± 0.0003	$(2.23 \pm 0.04) \times 10^{-5}$
	Low	2011+2012	0.1519 ± 0.0006	0.0212 ± 0.0007	0.0180 ± 0.0003	$(5.80 \pm 0.17) \times 10^{-5}$
	Low	2015+2016	0.1573 ± 0.0009	0.0151 ± 0.0005	0.0081 ± 0.0001	$(1.92 \pm 0.06) \times 10^{-5}$
	Low	2017+2018	0.1600 ± 0.0010	0.0205 ± 0.0008	0.0130 ± 0.0004	$(4.26 \pm 0.09) \times 10^{-5}$
	High	2011+2012	0.1509 ± 0.0011	0.0063 ± 0.0009	0.0093 ± 0.0006	$(8.90 \pm 1.10) \times 10^{-6}$
	High	2015+2016	0.1562 ± 0.0015	0.0056 ± 0.0003	0.0059 ± 0.0002	$(5.15 \pm 0.24) \times 10^{-6}$
	High	2017+2018	0.1576 ± 0.0016	0.0056 ± 0.0002	0.0072 ± 0.0002	$(6.38 \pm 0.19) \times 10^{-6}$

Table 7.6: Calculated total efficiencies for the rare mode processes for the different q^2 regimes. The quoted uncertainty is statistical. All values presented are calculated using candidates taken from the low $m(\pi^+\pi^-)$ bin and from the relevant q^2 bin. These values are used to determine the final rare mode branching fraction results.

CHAPTER 8

Invariant Mass Fits

“ In God we trust, all others must bring data ”

W. Edwards Deming,

The final branching fraction results are calculated using yields taken from unbinned, maximum likelihood fits to the $\pi^+\pi^-\mu^+\mu^-$ mass distribution. These mass fits make use of probability density functions (PDFs) which model the different signal and background contributions to the mass spectra. Before any fits are made to the data itself, the fitting models are tested on the simulated MC samples to determine the optimal PDF parameters for use in the final fits. These initial fits are completed over the same mass ranges as the final fits to ensure that the definition of the PDF parameters remains consistent.

Three types of invariant mass fits are performed within this analysis. The first set of fits are made to the $B^0 \to J/\psi K^{*0}$ control mode sample. The yields from these fits are used to scale the final branching fraction results and are also used throughout the analysis to help estimate various background yields. The second set of fits are made to the resonant mode decays whose purity compared to the rare modes makes them a useful proxy for testing the different PDF shapes and validating the broader fitting strategy. Finally, the rare mode fits themselves are performed. Unlike the case for the resonant and control mode fits, for the rare mode fits the 2011+2012, 2015+2016 and 2017+2018 datasets are fit simultaneously to extract the two branching fractions of interest at the same time. Before the rare mode fits are completed a series of pseudo-experiments are run to test the fitting strategy and look for any signs of bias which may affect the final result.

This chapter introduces the different invariant mass fits made within the analysis. For each fit the different PDF shapes which are used to model the signal and background components are introduced, the parameter values after the completion of the fits are tabulated and, finally, the fits to both MC and data samples are presented. Despite the initial fits to simulation being performed over the same mass ranges as the final fits, they are sometimes displayed within this chapter over a reduced range for visual clarity. Throughout this chapter reference is made to Appendix B which contains the fit parameters to the initial MC and data fits which are included for completeness. In addition, Appendix C contains the invariant mass fits from this chapter with the inclusion of additional pull plots which provide more detail on the quality of each fit.

8.1 Control Mode Fit

8.1.1 Control Mode Selection

The $B^0 \to J/\psi K^{*0}$ control mode fit is made separately for the three datasets and is selected by accepting candidates with $m(J/\psi K^+\pi^-) \in [5170, 6370]\,\mathrm{MeV}/c^2$ and $m(K^+\pi^-) \in [822, 962]\,\mathrm{MeV}/c^2$ where one of the final state pions is reconstructed with the known kaon mass. The $m(K^+\pi^-)$ requirement is specifically designed to select out $K^+\pi^-$ candidates originating from a $K^{*0}(892)$ resonance. This is a narrower window in $m(K^+\pi^-)$ mass than the one used for the kinematic weighting procedure discussed in Chapter 5. The selection here is designed to align this analysis with measurements of the S-wave fraction found in the previous literature [155] so that this fraction can be used in the calculation of the final branching fractions as was described in Chapter 4.

8.1.2 Control Mode Shapes

8.1.2.1 $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi K^{*0}$

The $B^0 \to J/\psi K^{*0}$ signal shape is described using the sum of two Crystal Ball (CB) functions [193] and a Gaussian function with the resulting PDF shown in Equation 8.1. The double CB + Gaussian shape is used throughout the analysis and is used to describe the signal components in the fits. The utility of the CB shape stems from the fact that the CB is a Gaussian with an additional power law tail. A CB shape with a power law tail at low invariant masses can be used to model the radiation emitted from final state particles. This radiation is often not reconstructed as part of the candidate with the effect being to smear out the lower end of the invariant mass distribution. In addition, a second CB shape with a power law tail at higher invariant mass values can model the candidates which are not well reconstructed and not well described with a simple

Gaussian resolution. The final Gaussian component is then included to account for the fact that the resolution of the reconstructed B meson is not constant over all candidates present within the dataset. The model is given by

$$\begin{aligned} S_{B^0\to J/\psi K^{*0}}(m) = f_{\rm CB}\cdot[f_1\cdot F_{\rm CB}(m;\mu_{B^0},\sigma_{\rm CB},\alpha_1,n_1) \\ + (1-f_1)\cdot F_{\rm CB}(m;\mu_{B^0},\sigma_{\rm CB},\alpha_2,n_2)] \\ + (1-f_{\rm CB})\cdot F_{\rm G}(m;\mu_{B^0},\sigma_{\rm G})\,. \end{aligned} \tag{8.1}$$

From Equation 8.1 it can be seen that the two CB functions share a common width, $\sigma_{\rm CB}$, and have tails at opposite ends of the mass distribution that are described by the different parameters, α_1, α_2 and n_1, n_2. The fraction of candidates that fall within the double CB shape, i.e. are not described by the Gaussian, is included in the factor $f_{\rm CB}$. Of those candidates not described by the Gaussian, the factor f_1 describes the fraction of these candidates described by the first CB. The Gaussian component then has a separate width $\sigma_{\rm G}$, but shares the same mean, μ_{B^0}, as the two CB shapes.

The values of these parameters are determined with initial fits to the $B^0\to J/\psi K^{*0}$ MC sample. For the final fit to data, the fractions $f_{\rm CB}$ and f_1, and the ratio $\sigma_{\rm G}/\sigma_{\rm CB}$ are fixed to the MC values. In addition, the final control mode data fit initialises the $\alpha_1, \alpha_2, \mu_{B^0}$ and $\sigma_{\rm CB}$ parameters to the values from the fits to simulation. However these parameters are allowed to float within the final fit as this is seen to result in an improved fit stability. In addition, the CB parameter, n, can become poorly defined at large values and it is possible to fix this parameter with no noticeable drop in performance. The parameter n_2 is therefore fixed to a value of 10 in the fit to MC while n_1 is allowed to float as before.

The PDF for the background process $B^0_s\to J/\psi K^{*0}$ shares the same shape and parameter values as the $B^0\to J/\psi K^{*0}$ mode. The only exception is that this background mode has a different mean, $\mu_{B^0_s}$, which is shifted compared to μ_{B^0} by the $87.4\,{\rm MeV}/c^2$ difference between the known B^0 and B^0_s masses [154].

The fits to the $B^0\to J/\psi K^{*0}$ MC samples for the three separate datasets is shown in Figure 8.1 where the fit is shown up to an invariant mass of $5600\,{\rm MeV}/c^2$ for visual clarity. As can be seen from the fits the signal is well described by the chosen PDF shape. The corresponding parameter values can be found in Table B.1.

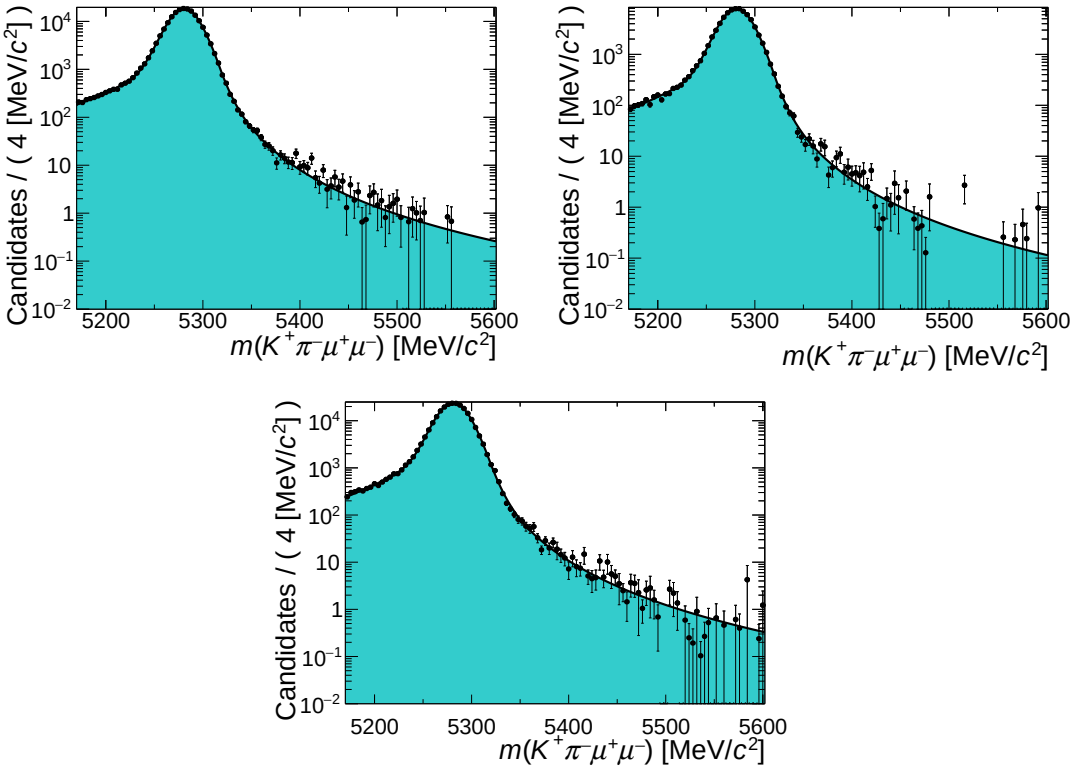

Figure 8.1: Invariant mass fits to the $B^0 \to J/\psi K^{*0}$ MC samples using the sum of two Crystal Ball functions and a Gaussian function for the 11+12 (top left), 15+16 (top right) and 17+18 (bottom) MC samples. The candidates are weighted with the full set of analysis correcting weights.

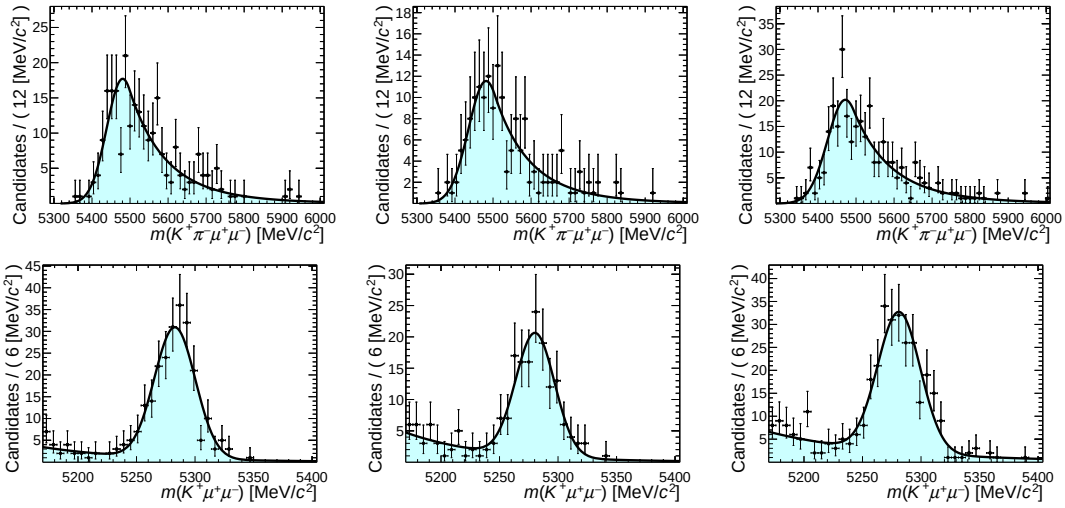

Figure 8.2: Invariant mass fits to the three datasets to determine the shape and yield of $B^+ \to J/\psi\, K^+$ decays present within the final control mode fit. The top row shows the fit to the $m(K^+\pi^-\mu^+\mu^-)$ invariant mass using a single CB shape for the 2011+2012 (left), 2015+2016 (middle) and 2017+2018 (right) datasets to extract a background shape. The bottom row shows, for the same three datasets, Gaussian fits to the $m(K^+\mu^+\mu^-)$ distribution to extract an estimated background yield. Both sets of fits include an exponential component to account for any remaining combinatorial background.

8.1.2.2 $B^+ \to J/\psi\, K^+$

As was discussed in Chapter 5, over-reconstructed background from $B^+ \to J/\psi\, K^+$ decays are expected to be present within the control data. This source of background must be included with a separate PDF component which takes the form of a single CB function. The parameter values in this instance are determined using data as it is possible to identify the background directly from the $K^+\mu^+\mu^-$ and $K^+\pi^-\mu^+\mu^-$ mass spectra. Furthermore, the background yield can be obtained from these fits to data and can then be scaled with PID and mass efficiencies and then imposed as a constraint on the yield in the final control mode fit. This constraint is applied within the final control mode fit as a Gaussian constraint.

The shape parameters for the background are determined by fitting a single CB shape to the $K^+\pi^-\mu^+\mu^-$ mass spectra where the K^+ is mis-identified as a π^+. Within this fit the CB parameter, n, is fixed to a value of 10 as previously. The yield is determined with a separate fit to the $K^+\mu^+\mu^-$ mass distribution where the peak is modelled with a single Gaussian shape. Both of these fits include an exponential component which is included to model any remaining combinatorial background present within the data. Both sets of fits can be seen in Figure 8.2 with the corresponding fit parameters in Table B.2.

8.1.2.3 Control Mode Total PDF

In the final control mode fits, a combinatorial background component, modelled with an exponential function, is included alongside the signal and peaking background components. The corresponding exponential parameter, λ, is allowed to float freely in the fit. Then, by combining the PDF shapes of the signal and background components, the final control mode PDF is written as is shown in Equation 8.2.

$$\begin{aligned} S(m) = \frac{1}{N_{\text{total}}} \Big(& N(B^0 \to K^{*0} J/\psi) \cdot S_{B^0 \to K^{*0} J/\psi}(m) \\ & + N(B_s^0 \to K^{*0} J/\psi) \cdot S_{B_s^0 \to K^{*0} J/\psi}(m) \\ & + N(B^+ \to J/\psi K^+) \cdot S_{B^+ \to J/\psi K^+}(m) \\ & + N_{\text{bkg.}} S_{\text{bkg}}(m) \Big) \, , \end{aligned} \tag{8.2}$$

Within this PDF it is assumed that each of the individual PDFs have been normalised over the $[5170-6370]\,\text{MeV}/c^2$ fit range and all of the prior fits to MC and data are made over this range for consistency. The final control mode fits can be seen in Figure 8.3 where the different components are overlapped in a stacked plot to show how the total fit model is built up using the individual components. In order to highlight the purity of the control mode, the fits can be seen again in Figure 8.4, on a linear scale. The final fit parameters from the control mode fits can be found in Table 8.1.

Parameter	2011+2012	2015+2016	2017+2018
α_1	-2.09 ± 0.04	-1.99 ± 0.04	-1.88 ± 0.03
α_2	0.45 ± 0.01	0.48 ± 0.00	0.52 ± 0.33
μ_{B^0}	5284.58 ± 0.05	5279.87 ± 0.05	5280.03 ± 0.04
σ_{CB}	19.88 ± 0.06	20.91 ± 0.06	20.85 ± 0.07
λ	-0.008 ± 0.001	-0.005 ± 0.002	-0.008 ± 0.006
$N(B^0 \to K^{*0} J/\psi)$	134244 ± 395	139344 ± 400	287350 ± 1707
N_{bkg}	666 ± 143	187 ± 136	1328 ± 1686
$N(B_s^0 \to K^{*0} J/\psi)$	1198 ± 62	1202 ± 62	2466 ± 129
$N(B^+ \to J/\psi K^+)$	210 ± 15	137 ± 12	226 ± 28

Table 8.1: Parameters from the fit to the $B^0 \to J/\psi K^{*0}$ control mode for the 2011+2012, 2015+2016 and 2017+2018 datasets.

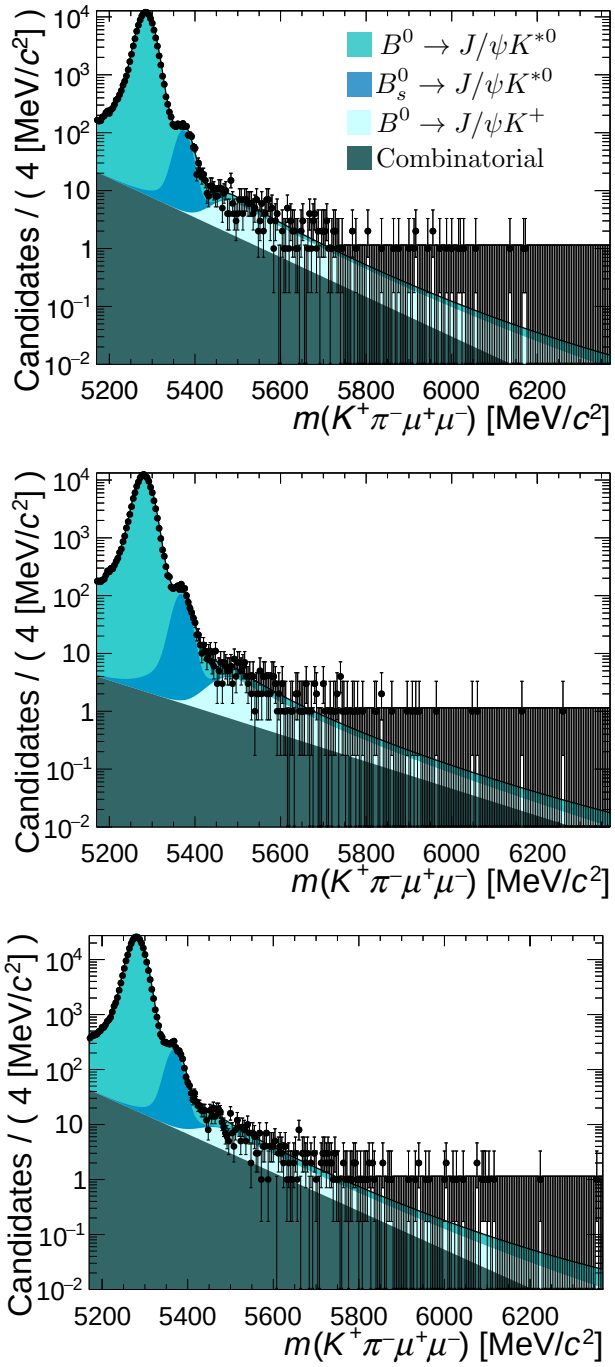

Figure 8.3: The invariant mass fits to the $B^0 \to J/\psi K^{*0}$ control mode for the 2011+2012 dataset (top), the 2015+2016 dataset (middle) and the 2017+2018 dataset (bottom).

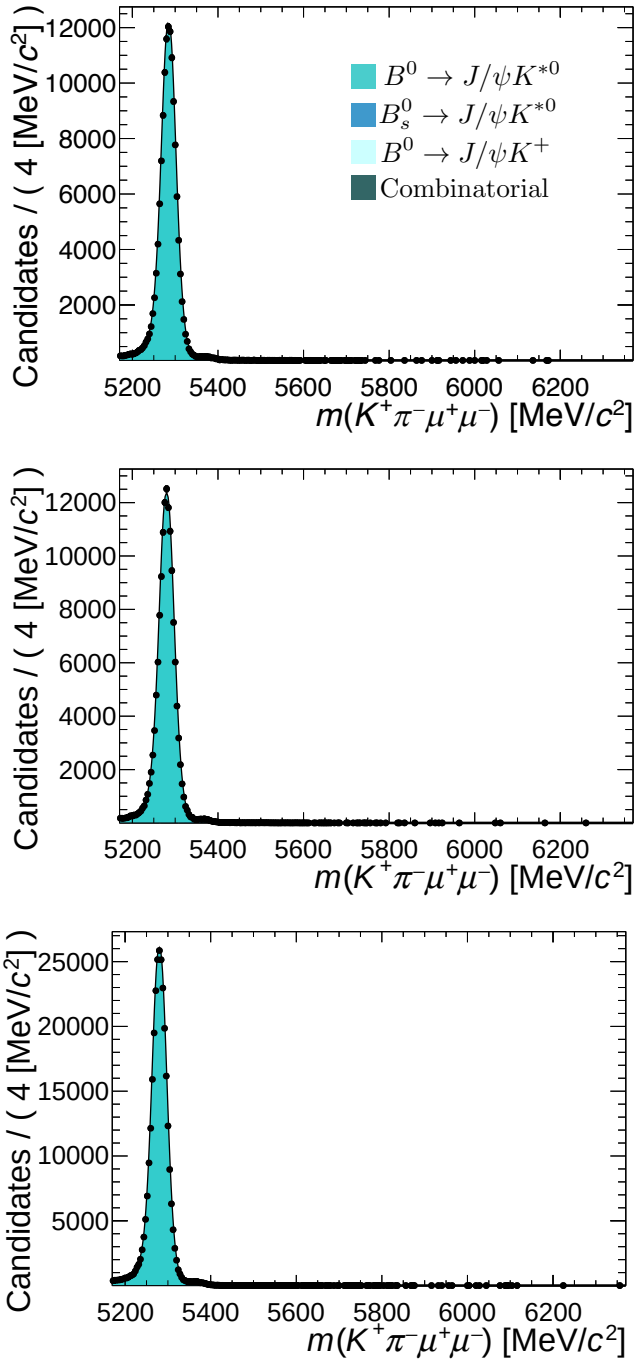

Figure 8.4: The invariant mass fits to the $B^0 \to J/\psi K^{*0}$ control mode for the 2011+2012 dataset (top), the 2015+2016 dataset (middle) and the 2017+2018 dataset (bottom). The linear scale makes clear the purity of the signal mode.

8.2 Resonant Mode Fit

8.2.1 Resonant Mode Fit Selection

The resonant mode fits are performed in a mass window corresponding to $m(J/\psi\,\pi^+\pi^-) \in [5170, 6370]\,\mathrm{MeV}/c^2$. The reason for the high upper boundary is to include the turnover of the B_c^+ background which was shown in Chapter 5 to be clearly visible at higher invariant mass values. The mass fits to the resonant modes are made separately for the 2011+2012, 2015+2016 and 2017+2018 datasets. In addition, for each dataset, two fits are performed; one with, and one without the reconstructed dimuon mass set to the known J/ψ mass value. Constraining the dimuon mass in this way reduces the width of the peak components as experimental mass resolution smearing is artificially reduced. The narrower signal peaks are easier to resolve from the background when applying this constraint. In addition, if the resonant mode fits are stable, then consistency should be seen between the resonant mode fits performed with and without a J/ψ mass constraint. It should be pointed out at this stage that this comparison cannot be made for the final rare mode fits due to the dimuon pair in the rare mode fits not originating from a narrow q^2 region which is synonymous with the J/ψ charmonium resonance.

8.2.2 Resonant Mode Signal shape

The $B^0 \to J/\psi\,\pi^+\pi^-$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ signal components are described by the sum of two CB functions and a Gaussian function, as was done for the control mode. The PDF parameters are determined from initial fits to the $B^0 \to J/\psi\,\pi^+\pi^-$ MC sample with the tail parameters n_1 and n_2 fixed to a value of 10.0 as before. Within the resonant mode fits the width of the Gaussian is not shared with the double CB shape. Instead, the ratio of the Gaussian and CB width is included as a parameter in the fit. Furthermore, within the final resonant mode fits, the mean and CB width parameter are initialised to the values determined from the MC fits however are allowed to float as this is found to result in an improved fit performance.

The fits to the signal MC sample can be seen in Figure 8.5 for both the case with and without the J/ψ mass constraint. The corresponding fit parameters can be found in Table B.3.

8.2.3 Resonant Mode Background shapes

The background $B^0 \to J/\psi\,K^+\pi^-$ resulting from particle mis-identification and the partially reconstructed backgrounds $B_s^0 \to J/\psi\,\eta'$ and $B^0 \to J/\psi\,\eta'$ are modelled with the sum of two CB functions which share a common width and mean. The fits to MC for the $B^0 \to J/\psi\,K^+\pi^-$ background fixes the n_2 value as before while the fit to the $B_s^0 \to J/\psi\,\eta'$ MC sample fixes both n

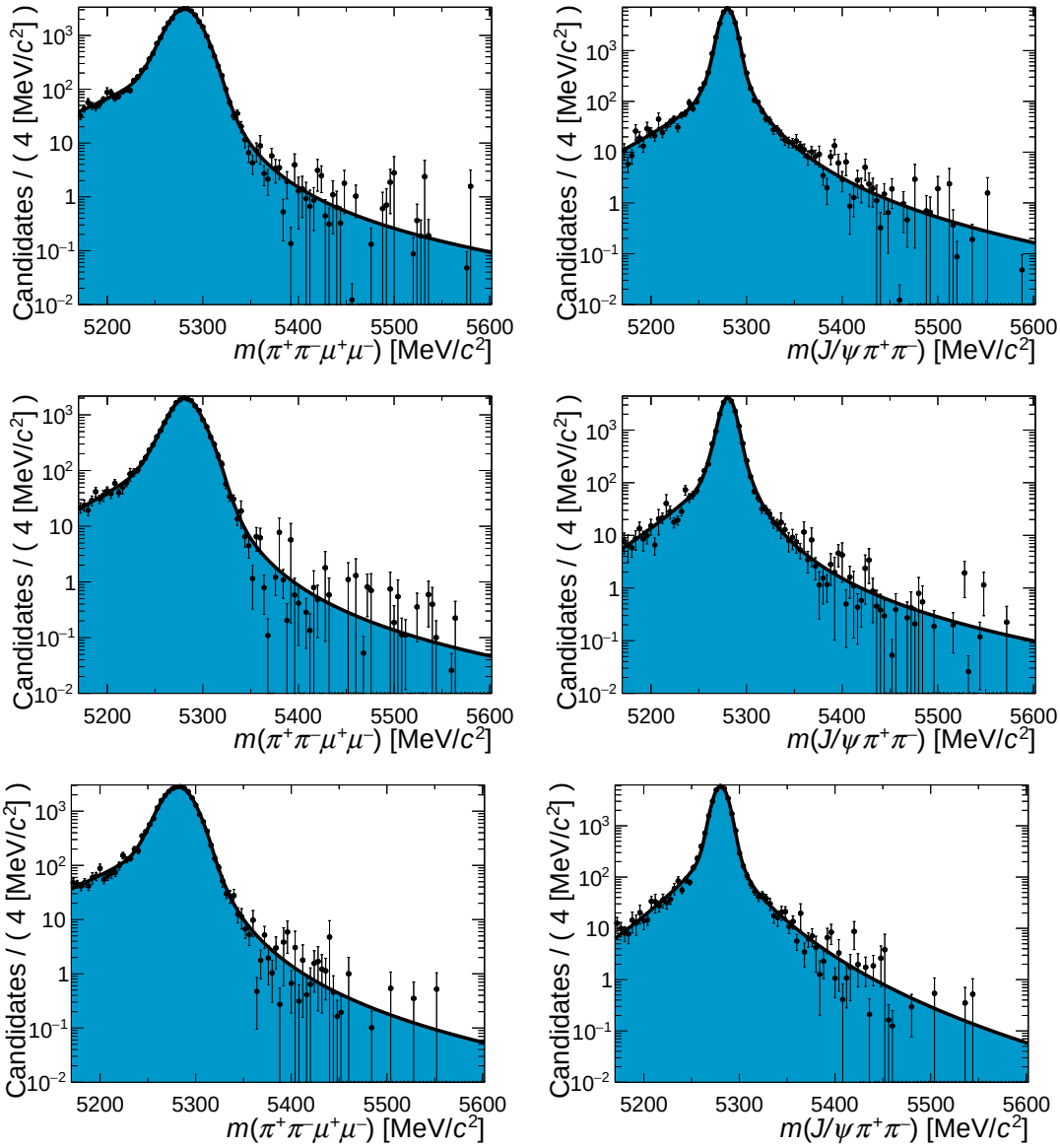

Figure 8.5: Invariant mass fits to the $B^0 \to J/\psi\pi^+\pi^-$ MC samples which are used to determine the signal shapes for the final resonant mode fits. The fits are performed with the sum of two CB shapes and a Gaussian shape. The left column, from top to bottom, shows the 2011+2012, 2015+2016 and 2017+2018 fits which are made without a constraint placed on the reconstructed J/ψ invariant mass. The right column shows the equivalent fits which include the constraint on the J/ψ mass. The candidates are weighted with the full set of analysis correcting weights.

parameters for improved fit stability. The $B^0 \to J/\psi\eta'$ background shares the same shape as the B_s^0 equivalent apart from the mean, which is shifted by the difference in the B^0 and B_s^0 masses.

The fits to the simulated samples for the $B^0 \to J/\psi K^+\pi^-$ and $B_s^0 \to J/\psi \eta'$ modes can be found in Figure 8.6 and Figure 8.7, where fits are shown, as before, with and without the J/ψ mass constraint. The parameters for these fits can be found in Table B.4 and Table B.5 respectively.

Some of the parameters from these MC fits are seen to have large associated uncertainties. This is due to the value of some of the correcting weights which are applied to the MC samples to improve the data-MC agreement being very small. This ultimately results in a very small effect ($< 1\%$) on the final resonant mode mass fit yields. As the resonant mass fits are made as a demonstrative precursor to the rare mode fits this is not considered problematic to the rest of the analysis. The possibility of improving the $B^0 \to J/\psi K^+\pi^-$ and $B_s^0 \to J/\psi \eta'$ mass fits was explored where the shapes were instead fit with a kernel density estimator (KDE) [194, 195]. The change in the yield of these components within the final resonant mode mass fits is smaller than the statistical uncertainty when using the KDE method however, and so the analysis continues to use the standard PDF shapes.

The $B_s^0 \to J/\psi \phi$ background, where $\phi \to \pi^+\pi^-\pi^0$, is modelled using a KDE. Using a KDE makes it easier to fit this background shape which is difficult to fit with a conventional lineshape. In addition, this background makes only a very small contribution towards the final mass fits and so using a KDE is appropriate. The KDE fit to the MC can be found in Figure 8.8.

A final background component is included to model the presence of B_c^+ decays to both the $J/\psi \pi^+\pi^-\pi^+$ and $J/\psi \pi^+\pi^-\pi^+\pi^+\pi^-$ final states. These two components are modelled with the `RapidSim` simulation package due to the unavailability of full LHCb MC simulation for these modes. The simulated samples, which decay the B_c^+ to the two possible final states, are then combined into a single sample in the ratio given by the branching fractions of $B_c^+ \to J/\psi \pi^+\pi^-\pi^+$ and $B_c^+ \to J/\psi \pi^+\pi^-\pi^+\pi^+\pi^-$ which are obtained from the Particle Data Group values [175]. When fitting the B_c^+ component, the mass spectrum is modelled with the combination of a bifurcated Gaussian and a standard Gaussian. The bifurcated Gaussian is a Gaussian where the width is different either side of the mean. The fit to the `RapidSim` sample can be seen in Figure 8.9 with the resulting fit parameters found in Table B.6.

Finally, as was done for the control mode, any remaining combinatorial background which survives the BDT selection requirement is parameterised with an exponential function.

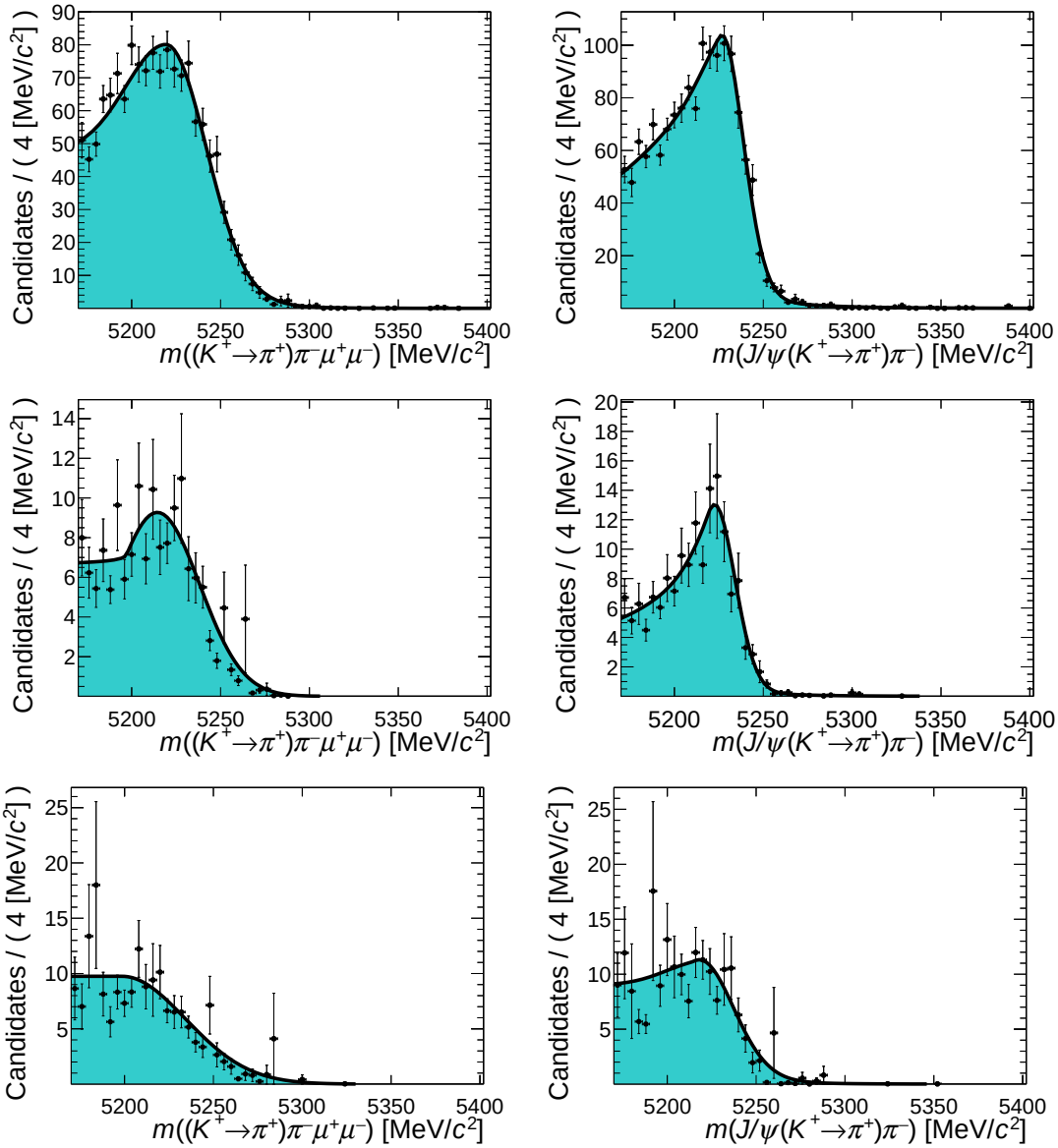

Figure 8.6: Invariant mass fits to the $B^0 \to J/\psi K^+\pi^-$ MC samples made with the sum of two CB shapes. The left column, from top to bottom, shows the 2011+2012, 2015+2016 and 2017+2018 fits without a constraint placed on the reconstructed J/ψ invariant mass, respectively. The right column shows the equivalent fits which include the constraint on the J/ψ mass. The candidates are weighted with the full set of analysis correcting weights.

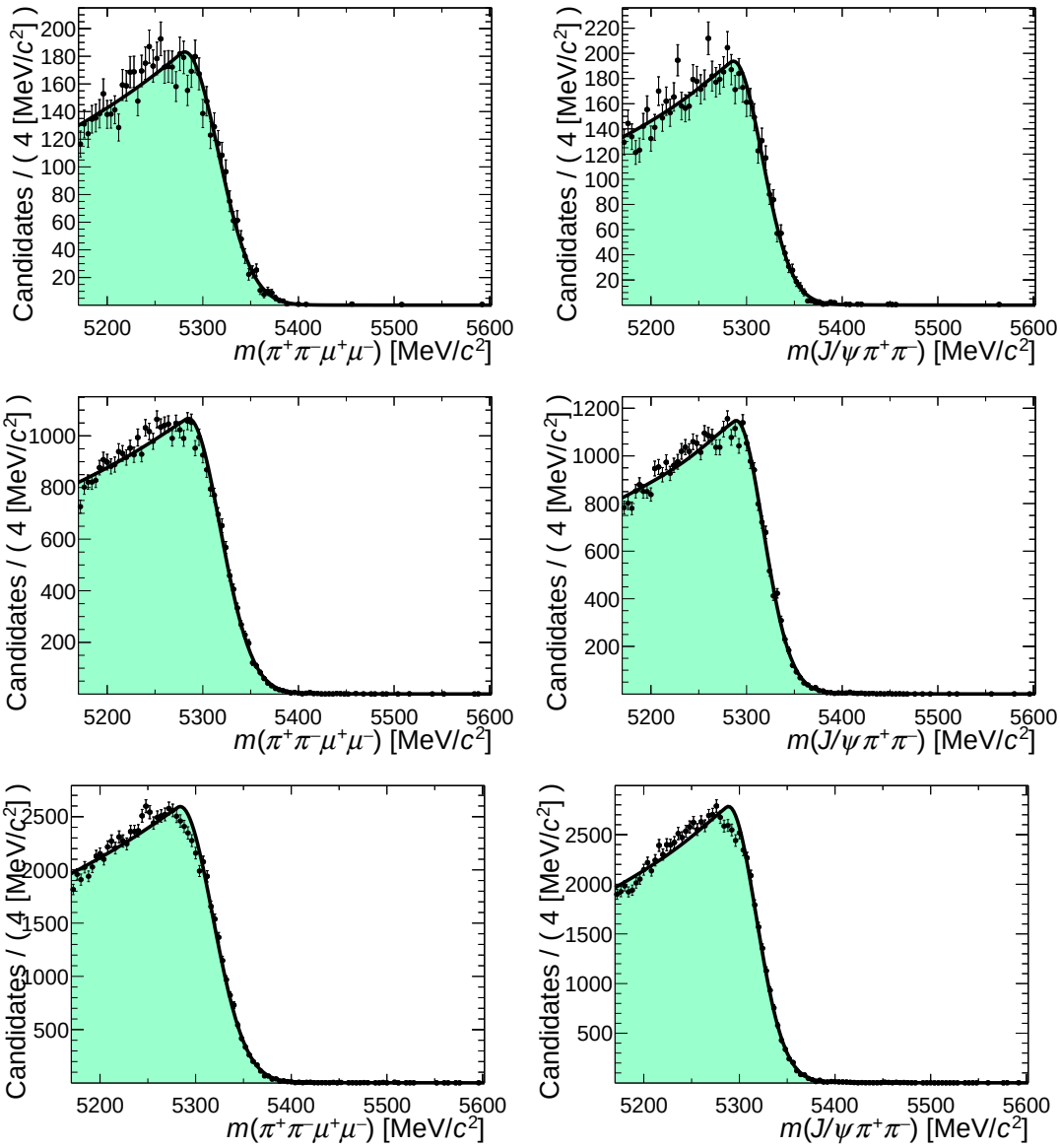

Figure 8.7: Invariant mass fits to the $B_s^0 \to J/\psi\eta'$ MC samples made with the sum of two CB shapes. The left column, from top to bottom, shows the 2011+2012, 2015+2016 and 2017+2018 fits without a constraint placed on the reconstructed J/ψ invariant mass, respectively. The right column shows the equivalent fits which include the constraint on the J/ψ mass. The candidates are weighted with the full set of analysis correcting weights.

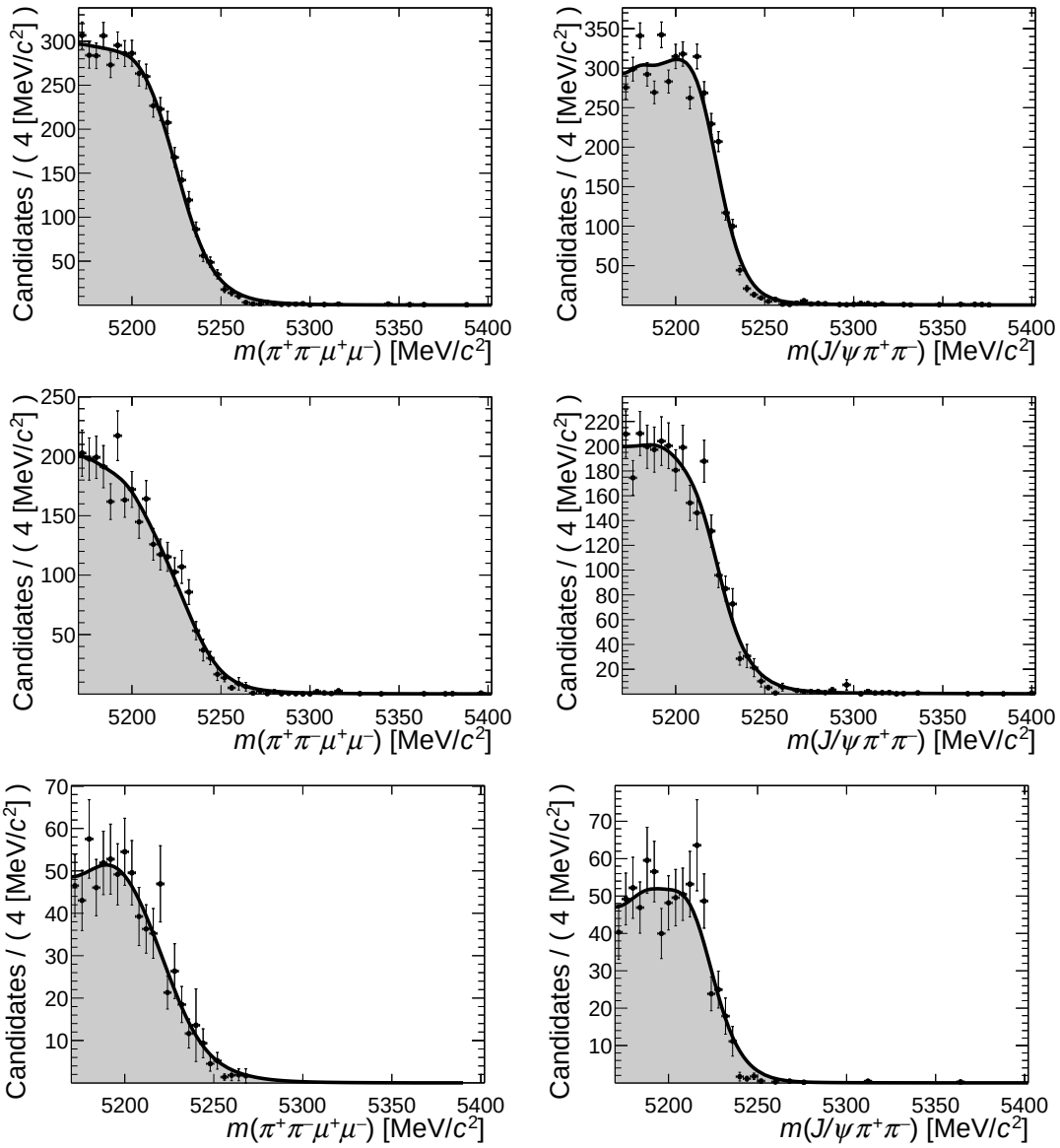

Figure 8.8: Invariant mass fits to the $B_s^0 \to J/\psi\,\phi$ MC samples implemented via a KDE. The left column, from top to bottom, shows the 2011+2012, 2015+2016 and 2017+2018 fits without a constraint placed on the reconstructed J/ψ invariant mass respectively. The right column shows the equivalent fits which include the constraint on the J/ψ mass. The candidates are weighted with the full set of analysis correcting weights.

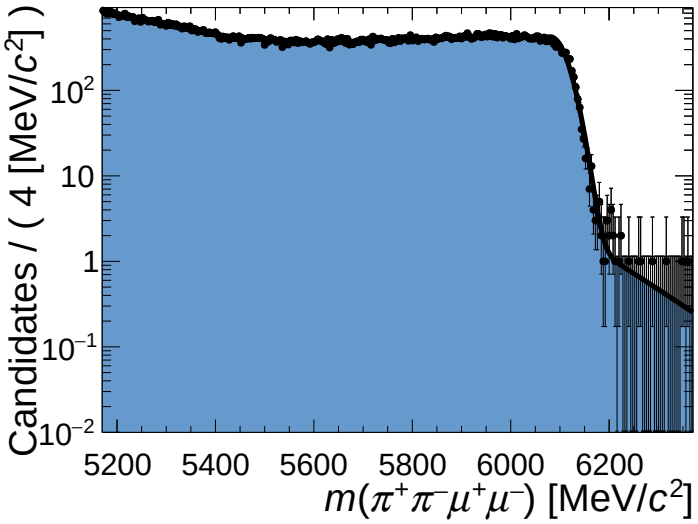

Figure 8.9: The invariant mass fit to the `RapidSim` simulated sample of B_c^+ background.

When performing the resonant mode fits the different signal and background components are combined to create the final PDF shown in Equation 8.3 where each of the individual PDF components have been normalised over the same mass range.

$$
\begin{aligned}
S(m) = \frac{1}{N_{\text{total}}} \Big(& N(B^0 \to J/\psi \pi^+\pi^-) \cdot S_{B^0 \to J/\psi \pi^+\pi^-}(m) + \\
& N(B^0_s \to J/\psi \pi^+\pi^-) \cdot S_{B^0_s \to J/\psi \pi^+\pi^-}(m) + \\
& N(B^0 \to J/\psi K^+\pi^-) \cdot S_{B^0 \to J/\psi K^+\pi^-}(m) + \\
& N(B^0_s \to J/\psi \eta') \cdot S_{B^0_s \to J/\psi \eta'}(m) + \\
& N(B^0_s \to J/\psi \phi) \cdot S_{B^0_s \to J/\psi \phi}(m) + \\
& N(B^+_c \to J/\psi 3\pi(\pi\pi)) \cdot S_{B^+_c \to J/\psi 3\pi(\pi\pi)}(m) + \\
& N_{\text{bkg}} S_{\text{bkg}}(m) \Big) \, ,
\end{aligned}
\tag{8.3}
$$

Within the final fit, the different background sources that lie in the low invariant mass region compete with each other. This means that the fitting procedure may converge on several different local minima where the yields of these different backgrounds are interchanged. In order to help the fitting procedure converge on a consistent minima the yields of the $B^0 \to J/\psi K^+\pi^-$, $B^0_s(B^0) \to J/\psi \eta'$, and $B^0_s \to J/\psi \phi$ backgrounds are each constrained.

The $B^0 \to J/\psi K^+\pi^-$ yield is constrained with a Gaussian function where the mean is set to the control mode fit yield scaled to account for the mass, BDT and PID efficiencies under the $K^+ \to \pi^+$ hypothesis relevant for the resonant mode fit. It is possible to write this scaled yield as is shown in Equation 8.4.

$$
N(B^0 \to J/\psi K^+\pi^-) = \tilde{N}(B^0 \to J/\psi K^+\pi^-) \times \frac{\epsilon(B^0 \to J/\psi K^+\pi^- | K \to \pi)}{\epsilon(B^0 \to J/\psi K^+\pi^- | K \to K)} \, . \tag{8.4}
$$

In Equation 8.4 $\tilde{N}(B^0 \to J/\psi K^+\pi^-)$ is the yield of $B^0 \to J/\psi K^+\pi^-$ decays from the fit to the $B^0 \to J/\psi K^{*0}$ control mode. The width of the Gaussian constraint is taken from the uncertainty on the control mode signal yield, propagated with the uncertainties on the efficiency ratio where the efficiency uncertainties are taken from the bootstrapping method discussed in Chapter 7.

The yields for the $B^0_s \to J/\psi \eta'$, $B^0 \to J/\psi \eta'$ and $B^0_s \to J/\psi \phi$ modes are constrained relative to the $B^0_s \to J/\psi \pi^+\pi^-$ yield. This constraint is implemented by scaling the yield of the B^0_s mode with ratios of efficiencies and fragmentation fractions following Equation 8.5.

$$
N(X) = N(B^0_s \to J/\psi \pi^+\pi^-) \times \frac{\epsilon(X)}{\epsilon(B^0_s \to J/\psi \pi^+\pi^-)} \times \frac{f_q}{f_s} \, , \tag{8.5}
$$

Within Equation 8.5 q refers to either an s quark or a d quark as is appropriate. Due to the lack of full LHCb MC for the $B^0 \to J/\psi\eta'$ mode, the efficiencies for this mode are assumed to be the same as for the $B_s^0 \to J/\psi\eta'$ mode. The exception is the $m(\pi^+\pi^-\mu^+\mu^-) \in [5170, 6370]\,\mathrm{MeV}/c^2$ and $m(\pi^+\pi^-) \in [515, 1300]\,\mathrm{MeV}/c^2$ mass window efficiency which is estimated from a separate `RapidSim` sample.

The final resonant mode fits to the data can be seen in Figure 8.10 on a logarithmic scale, where the different components can be seen, and in Figure 8.11 on a linear scale, where the purity of the two signal peaks is visible. The fit parameters are given in Table 8.3. Some of the resonant mode fits show a small level of disagreement with the data in the region attributable to the various peaking backgrounds at low invariant mass. This is caused by the constraints placed upon the yields which are calculated using the full set of correcting weights some of which, as has been already shown, have large values. The small discrepancies in the low invariant mass region disappear when removing these constraints as can be seen in Figure 8.12 and Figure 8.13. The small level of disagreement makes negligibly small difference to the signal peaks within the fits.

Overall, the fits to the resonant mode are seen to be of good quality and consistent between the fits with and without the J/ψ mass constraint. It can be seen from the Run 1 fit that the value of the mean of the signal peak converges to a value slightly higher than the known B^0 mass. This is understood to be due to the lack of momentum scale correction which accounts for small deviations in the LHCb magnetic field experienced by each particle track. These corrections are not present within the LHCb dataset for Run 1 by default however are present by default within the Run 2 data.

Although not the focus of this thesis, the resonant mode fits can be validated by calculating branching fraction values from the fit yields for the B^0 and B_s^0 for the modes and comparing them to the known world values. The branching fractions are calculated by scaling the yield of the resonant mode fits similarly to as is done for the rare mode as was discussed in Chapter 4. The values of the resonant mode branching fractions for each dataset, alongside the previous LHCb values, are shown in Table 8.2.

Within Table 8.2 two alternative values are presented for the B_s^0 mode branching fraction. The first result uses the fragmentation fraction values taken from a previous LHCb analysis[138]. The second result, labelled as "Alternate" uses a second fragmentation fraction measurement taken from a separate LHCb analysis [157]. The agreement between the 2011+2012, 2015+2016 and 2017+2018 datasets when using the alternate value in the final calculation.

When comparing the resonant mode branching fraction results found here to those from the previous analysis it is necessary to consider any correlations between the uncertainties. Both the statistical and systematic uncertainties from this analysis are believed to be largely

uncorrelated with those from the previous LHCb analyses. The selection used for the Run 1 data here is different to that used in the previous analysis meaning that the Run 1 datasets are different. Furthermore, correlations between systematic uncertainties are not of concern primarily due to the different control modes used in this analysis and the previous analyses. When disregarding the possibility of large correlations between the analyses a good level of agreement between the values of the branching fractions from this analysis and those from the previous LHCb analyses can be seen.

Table 8.2: Branching fractions of $B^0 \to J/\psi\pi^+\pi^-$ and $B_s^0 \to J/\psi\pi^+\pi^-$ calculated using the yields of the resonant mode fits. The branching fractions for the B_s^0 mode is provided for the two possible Run 2 fragmentation fraction values.

Dataset	$B(B^0 \to J/\psi\pi^+\pi^-)[10^{-5}]$	$B(B_s^0 \to J/\psi\pi^+\pi^-)[10^{-4}]$
2011+2012	$3.78 \pm 0.08(\text{stat}) \pm 0.27(\text{syst})$	$1.85 \pm 0.03(\text{stat}) \pm 0.17(\text{syst})$
2015+2016	$3.88 \pm 0.10(\text{stat}) \pm 0.28(\text{syst})$	$2.03 \pm 0.04(\text{stat}) \pm 0.18(\text{syst})$
2017+2018	$3.90 \pm 0.10(\text{stat}) \pm 0.28(\text{syst})$	$2.06 \pm 0.04(\text{stat}) \pm 0.18(\text{syst})$
2015+2016 (alternative)	–	$1.79 \pm 0.04(\text{stat}) \pm 0.17(\text{syst})$
2017+2018 (alternative)	–	$1.82 \pm 0.04(\text{stat}) \pm 0.17(\text{syst})$
Previous LHCb value	$3.94 \pm 0.14(\text{stat}) \pm 0.21(\text{syst})$	$2.09 \pm 0.08(\text{stat}) \pm 0.21(\text{syst})$

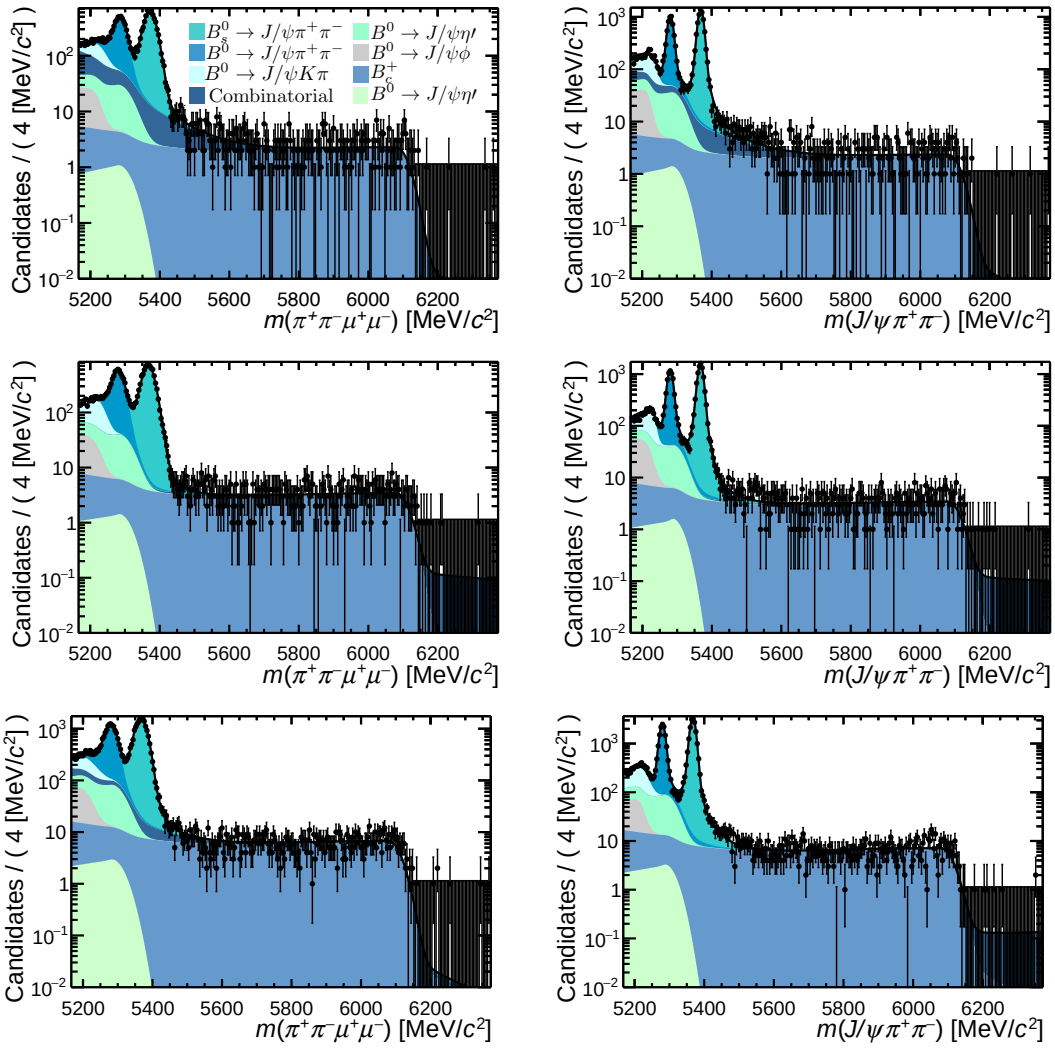

Figure 8.10: The final resonant mode invariant mass fits for the 2011+2012 (top), 2015+2016 (middle) and 2017+2018 (bottom) datasets on a logarithmic scale to show the different background components. The left column shows the fits without the J/ψ mass constraint applied and the right plots with the J/ψ mass constraint applied.

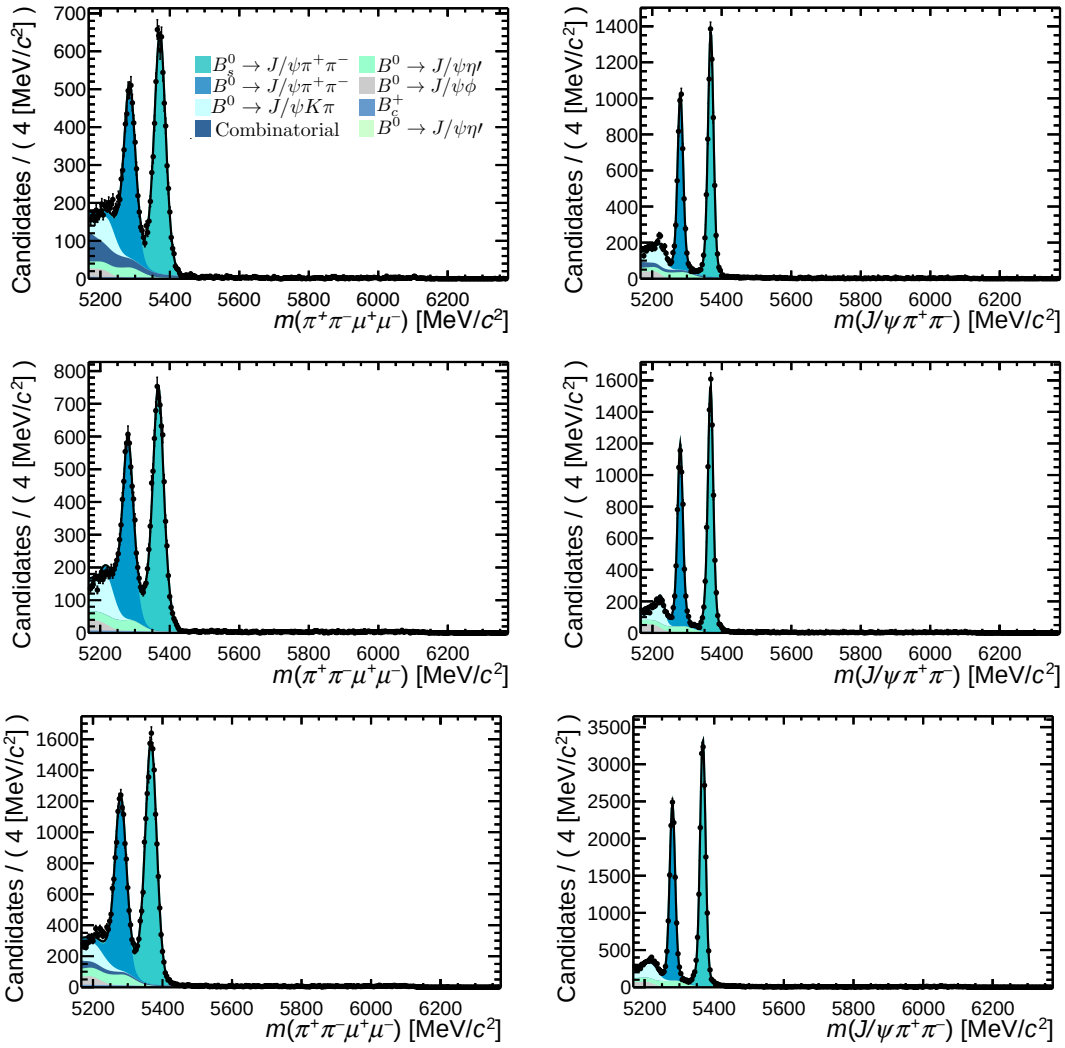

Figure 8.11: The final resonant mode invariant mass fits for the 2011+2012 (top), 2015+2016 (middle) and 2017+2018 (bottom) datasets. The left column shows the fits without the J/ψ mass constraint applied and the right plots with the J/ψ mass constraint applied.

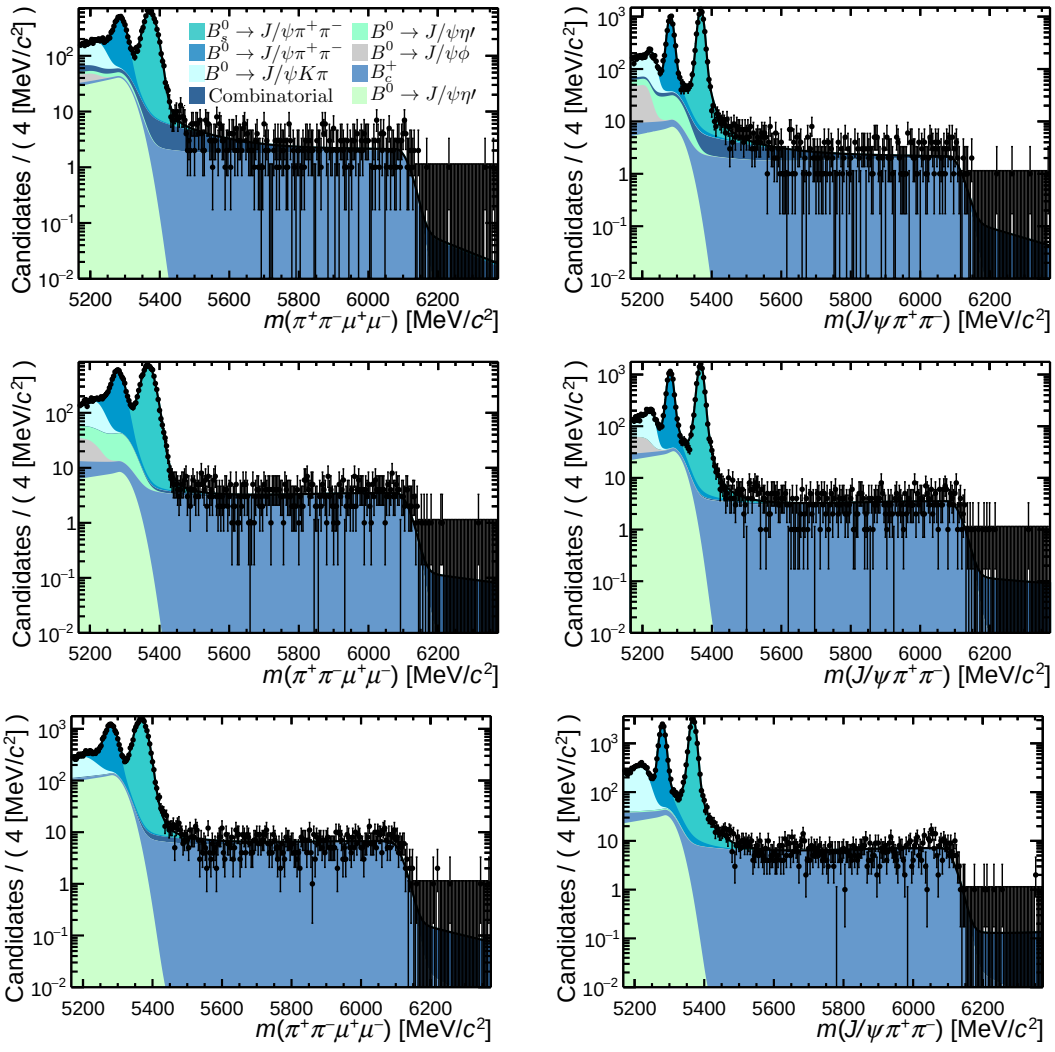

Figure 8.12: The final resonant mode invariant mass fits for the 2011+2012 (top), 2015+2016 (middle) and 2017+2018 (bottom) datasets on a logarithmic scale to show the different background components where no yield constraints are applied. The left column shows the fits without the J/ψ mass constraint applied and the right plots with the J/ψ mass constraint applied.

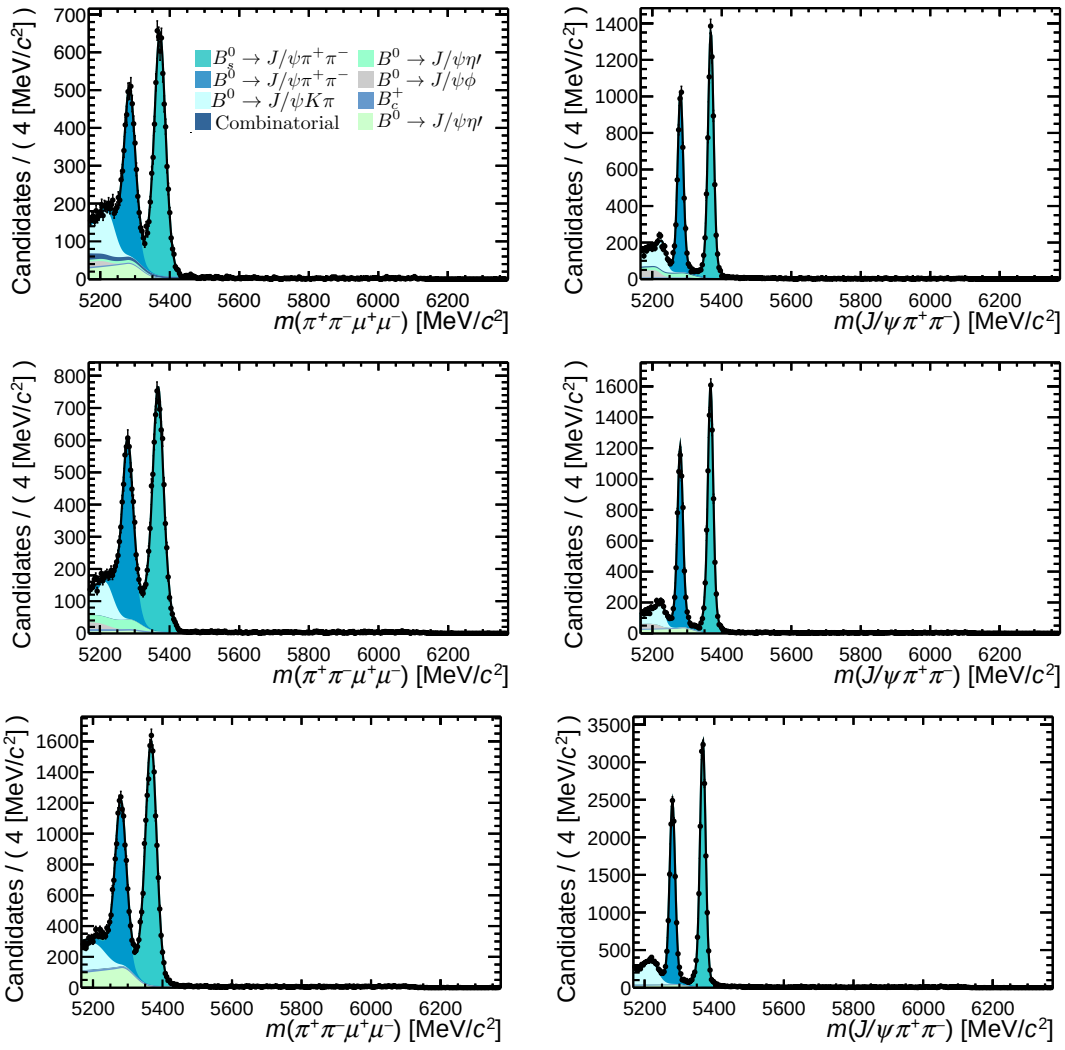

Figure 8.13: The final resonant mode invariant mass fits for the 2011+2012 (top), 2015+2016 (middle) and 2017+2018 (bottom) datasets where no yield constraints are applied. The left column shows the fits without the J/ψ mass constraint applied and the right plots with the J/ψ mass constraint applied.

Table 8.3: Parameters from the fit to the resonant mode for 2011+2012, 2015+2016 and 2017+2018 both with and without the J/ψ mass constraint.

Fit	Parameter	2011+2012	2015+2016	2017+2018
No J/ψ constraint	$N(B^0 \to J/\psi\pi^+\pi^-)$	5082 ± 92	6126 ± 92	12354 ± 141
	$N(B_s^0 \to J/\psi\pi^+\pi^-)$	7280 ± 94	8582 ± 95	17893 ± 146
	$N(B^0 \to J/\psi K^+\pi^-)$	1279 ± 37	2082 ± 62	2287 ± 93
	$N(B_c^+)$	574 ± 39	847 ± 53	1696 ± 65
	N_{bkg}	1857 ± 126	49 ± 48	1265 ± 204
	μ_{B^0}	5284.07 ± 0.21	5279.69 ± 0.19	5279.41 ± 0.13
	σ_{CB}	19.87 ± 0.21	20.47 ± 0.19	18.99 ± 0.13
	λ	-0.010 ± 0.001	-0.001 ± 0.002	-0.010 ± 0.001
J/ψ constraint	$N(B^0 \to J/\psi\pi^+\pi^-)$	5333 ± 82	6327 ± 86	12863 ± 122
	$N(B_s^0 \to J/\psi\pi^+\pi^-)$	7460 ± 91	8624 ± 92	18031 ± 134
	$N(B^0 \to J/\psi K^+\pi^-)$	1654 ± 64	1968 ± 57	3317 ± 66
	$N(B_c^+)$	587 ± 50	849 ± 48	1781 ± 75
	N_{bkg}	885 ± 120	37 ± 35	24 ± 50
	μ_{B^0}	5281.23 ± 0.09	5279.89 ± 0.08	5279.84 ± 0.05
	σ_{CB}	14.07 ± 0.14	11.56 ± 0.10	12.61 ± 0.07
	λ	-0.009 ± 0.001	-0.000 ± 0.002	0.001 ± 0.005

8.3 Rare Mode Fit

In order to determine the branching fraction of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decays, the analysis uses a simultaneous fit to the 2011+2012, 2015+2016 and 2017+2018 datasets where the B^0 and B_s^0 branching fractions are freely varying as shared parameters between the datasets. Equation 4.2 from Chapter 4 can be re-expressed as:

$$\mathcal{B}(B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-) = N(B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-) \times \kappa \tag{8.6}$$

where, in Equation 8.6, κ, is referred to as the single event sensitivity and includes the control mode fit yield, rare mode and control mode total efficiency and fragmentation fraction information. Within the rare mode fits, both the B^0 and B_s^0 κ factors are allowed to vary within their uncertainties via the application of a Gaussian constraint.

To select out the rare mode datasets the full set of requirements from Table 5.6 are applied in addition to the stripping, trigger and PID fiducial requirements, as before. The main difference between the resonant mode and rare mode selection requirements, aside from the J/ψ mass veto, is the upper limit on the fitting region. Within the rare mode fits the upper limit is set to $5800\,\mathrm{MeV}/c^2$ as candidates above this were used for the definition of the upper sideband.

As was done for the resonant mode fits, the signal shapes are modelled with the sum of two CB functions and a Gaussian function while the $K^+ \to \pi^+$ mis-identified background from $B^0 \to K^+\pi^-\mu^+\mu^-$ decays is modelled with a pair of CB functions. The yield for this background is constrained using a Gaussian function. Due to the unknown S-wave composition of the $K^+\pi^-$ system in the rare mode fit, a separate double CB fit to the $K^+\pi^-\mu^+\mu^-$ invariant mass is made to the data where reversed PID requirements are made in order to extract a clear signal peak. The yield from this fit is then scaled with both mass and PID efficiencies in order to extrapolate the number of expected $K^+\pi^-\mu^+\mu^-$ background candidates to the rare mode fit. The resulting expected yield and yield uncertainty are used to set the mean and width of the Gaussian constraint respectively. These initial fits to the $K^+\pi^-\mu^+\mu^-$ invariant mass spectra are seen in Figure 8.14. The fit parameters from these fits can be found in Table B.7.

Initial fits are made to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ MC samples to determine the PDF parameters for the final fits. However, for the signal modes, only the tail parameters are set to the values from the MC mass fits. Both the mean and width are instead set to the values seen from the resonant mode fits as the shape parameters are expected to be the same between rare and resonant mode decays. Any variation in these shapes is expected to be small between q^2 regions and, given the very small number of expected candidates, can be disregarded. The fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC samples can be found in Figure 8.15 while

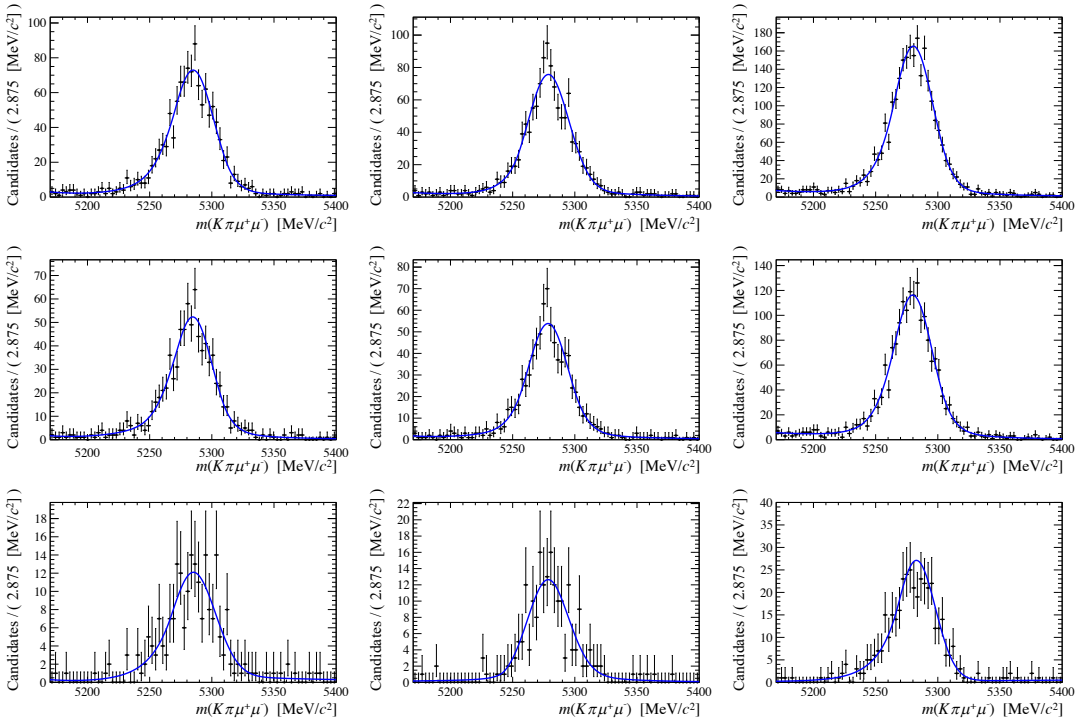

Figure 8.14: Fit to the $m(K^+\pi^-\mu^+\mu^-)$ distribution for 2011+2012 (left), 2015+2016 (middle) and 2017+2018 (right). The top row shows the fits to the full q^2 range, the middle row the fits to the low q^2 range and the bottom row the fits to the high q^2 range. The data is fit with the sum of two CB functions and an exponential. The yield is scaled using efficiencies in order to provide a constraint for the $B^0 \to K^+\pi^-\mu^+\mu^-$ background in the final rare mode fits.

the equivalent fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC can be found in Figure 8.16. The fit parameters for the signal mode MC can be found in Table B.8 while the fit parameters for the background mode MC can be found in Table B.9.

Finally, as was the case for both the control and resonant modes, a combinatorial background is included and is modelled with an exponential function. The final PDF fit to the rare mode invariant mass fits is given by

$$\begin{aligned} S(m) = \frac{1}{N_{\text{total}}}\Big(& N(B^0 \to \pi^+\pi^-\mu^+\mu^-) \cdot S_{B^0 \to \pi^+\pi^-\mu^+\mu^-}(m) + \\ & N(B_s^0 \to \pi^+\pi^-\mu^+\mu^-) \cdot S_{B_s^0 \to \pi^+\pi^-\mu^+\mu^-}(m) + \\ & N(B^0 \to K^+\pi^-\mu\mu^-) \cdot S_{B^0 \to K^+\pi^-\mu^+\mu^-}(m) + \\ & N_{\text{bkg}} S_{\text{bkg}}(m)\Big). \end{aligned} \tag{8.7}$$

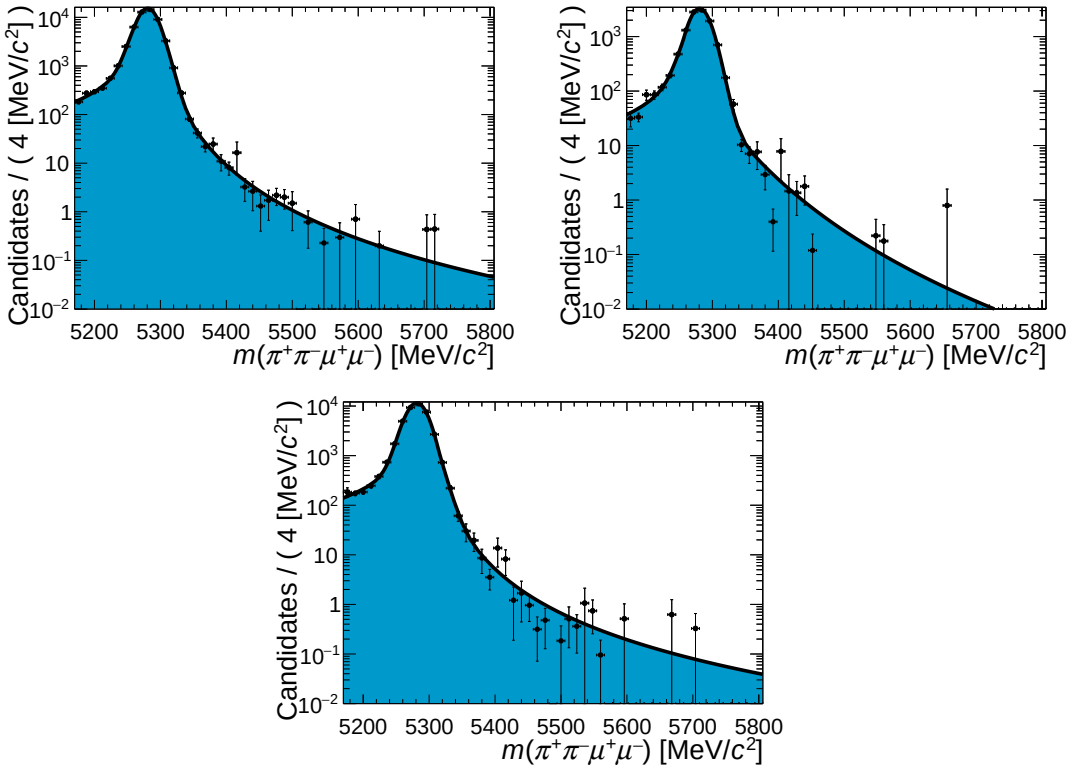

Figure 8.15: Invariant mass fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC samples which serves as a proxy for the signal shapes within the rare mode fits. The fits are performed with the sum of two CB shapes and a Gaussian shape and can be seen for the 2011+2012 sample (top left), the 2015+2016 sample (top right) and the 2017+2018 sample (bottom). The candidates are weighted with the full set of analysis correcting weights.

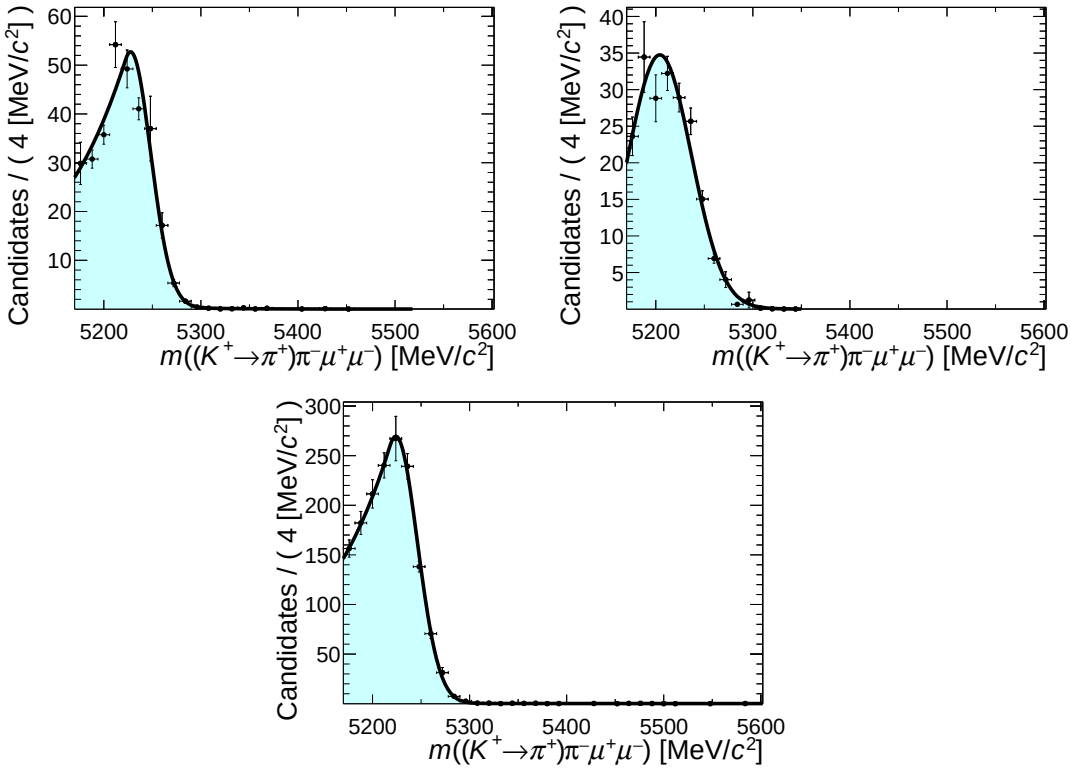

Figure 8.16: The invariant mass fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC samples for the 2011+2012 (top left), 2015+2016 (top right) and the 2017+2018 sample (bottom). The fits are performed using the sum of two CB shapes. The candidates are weighted with the full set of analysis correcting weights.

8.3.1 Fit Model Validation

Prior to the completion of the final rare mode fits a series of 100,000 pseudo-experiments are performed where data is generated using the model from Equation 8.7. These pseudo-experiments are run to validate the fitting strategy and detect any systematic bias in the model. The number of generated events is randomised per pseudo-experiment using a Poisson distribution with the mean set to the expected yield. The expected yields are calculated using Equation 4.2, where the rare mode branching fractions are assumed to be equal to the values determined by the Run 1 LHCb analysis. After the generation of the pseudo-experiment data, the total rare mode PDF is fit to the data and the branching fractions of the B^0 and B^0_s decay extracted. An example of a fit to the pseudo-experiment data can be seen in Figure 8.17. The branching fraction pull distributions, which is the residual of the branching fractions when compared to the previous values normalised by the corresponding uncertainty, can be seen in Figure 8.18 for both rare modes and for the three q^2 regimes.

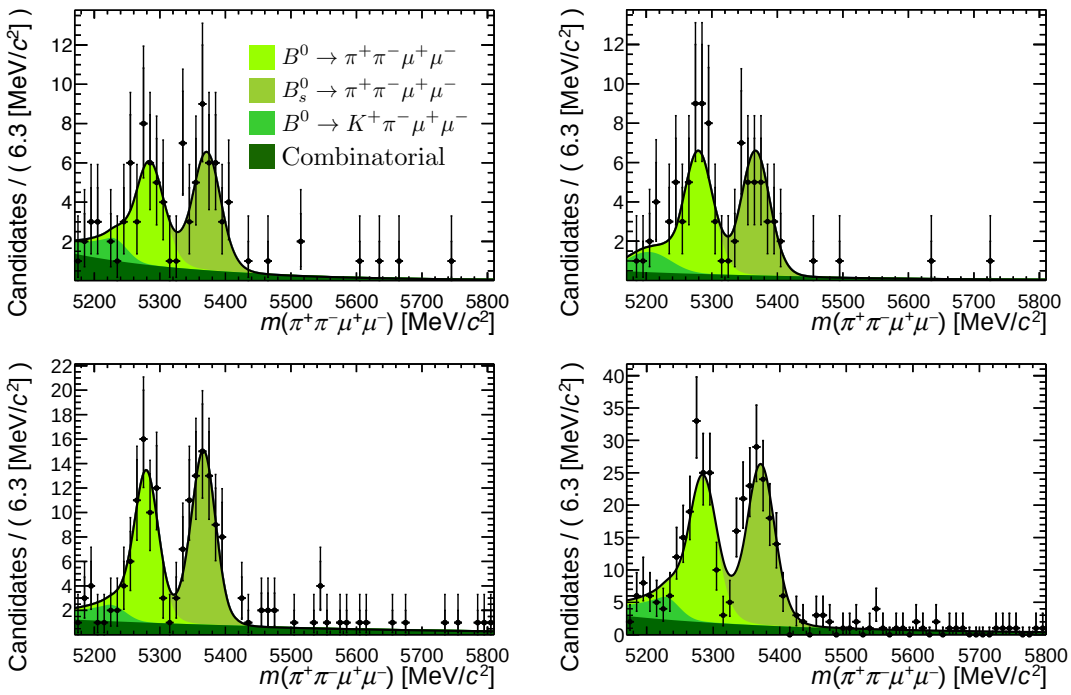

Figure 8.17: Examples of a simultaneous fit to the rare mode pseudo-experiment. Top left: fit to the 2011+2012 generated data. Top right: fit to the 2015+2016 generated data. Bottom left: fit to the 2017+2018 generated data. Bottom right: combined fit from the three simultaneous fits, included for visualisation.

Table 8.4: Results from the pseudo-experiment studies where the pull distributions are fit with a Gaussian.

q^2 Range	Decay	Mean	Standard Deviation
Full	B^0 Mode	-0.046 ± 0.003	0.988 ± 0.002
	B_s^0 Mode	-0.036 ± 0.003	0.989 ± 0.002
Low	B^0 Mode	-0.062 ± 0.003	0.988 ± 0.002
	B_s^0 Mode	-0.038 ± 0.003	0.983 ± 0.002
High	B^0 Mode	-0.070 ± 0.003	0.969 ± 0.002
	B_s^0 Mode	-0.092 ± 0.003	0.966 ± 0.002

The pull distributions are fit with Gaussian distributions. As can be seen from the results in Table 8.4 the means are broadly consistent with zero, while the widths are close to unity, as would be expected for unbiased fits. However, a small level of bias is seen, which is consistent across the different sets of pseudo-experiments, suggesting that the fit model slightly underestimates the branching fractions. This bias is caused by the small yields of the different fit components including the background component. Although this effect is very small (a bias of $\mathcal{O}(10^{-10})$ on a branching fraction value of $\mathcal{O}(10^{-8})$), the final branching fraction results are corrected for this bias. The correction is applied by shifting the final branching fraction results by the product of the final statistical uncertainty and the mean shift seen from Table 8.4. For example, the final shift for the B^0 mode for the full q^2 range is given by 4.6% multiplied by the final statistical uncertainty.

8.3.2 Final Rare Mode Fits

The result of the simultaneous invariant mass fits for the full q^2, low q^2 and high q^2 regimes can be found in Figures 8.19, 8.20 and 8.21 respectively. In each of these figures, the fits to the three datasets are shown and then combined into a fourth figure for visualisation. The final PDF parameters for the rare mode fits are found in Table 8.5. Within Table 8.5 the yields shown in blue are not included directly within the fits but are instead derived afterwards using the final branching fraction and event sensitivity parameters. These are included in the tables for useful comparison. In addition, the fits are repeated for all three q^2 regimes for each of the three datasets separately, without performing a simultaneous fit. This provides a consistency check for the analysis. Figure 8.22 shows the results of the fits to the individual datasets, while Table 8.6 shows the corresponding fit parameters. The results of the rare mode fits are the subject of Chapter 10.

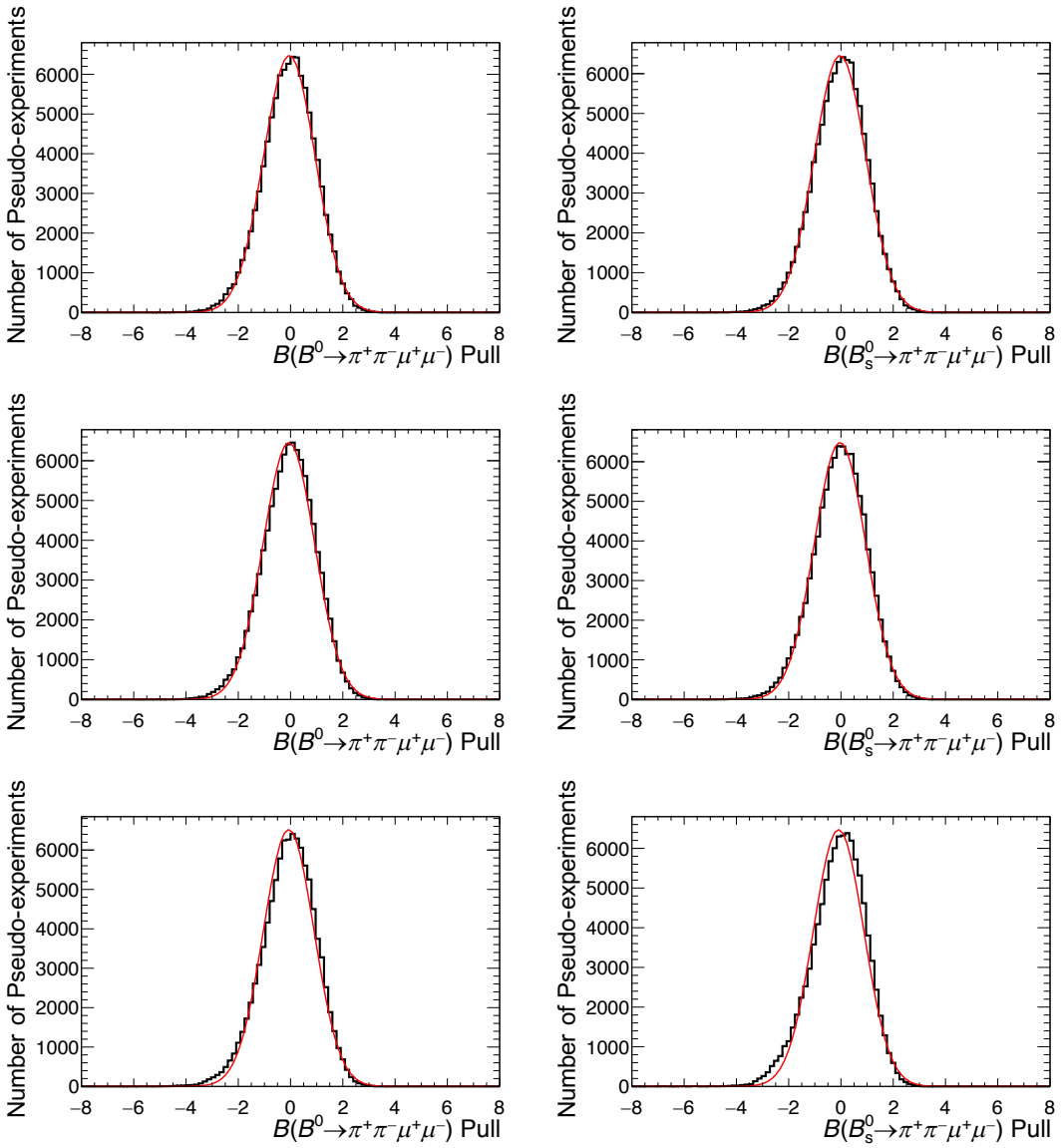

Figure 8.18: The distribution of the branching fraction pull (residual normalised by the branching fraction uncertainty) for the B^0 mode (left column) the B_s^0 mode (right column) from the pseudo-experiment study. The top row are the results for the full q^2 region while the middle row shows the results for the low q^2 region and the bottom row the results for the high q^2 region. The distributions are fit with a Gaussian to determine the presence of any fitting bias.

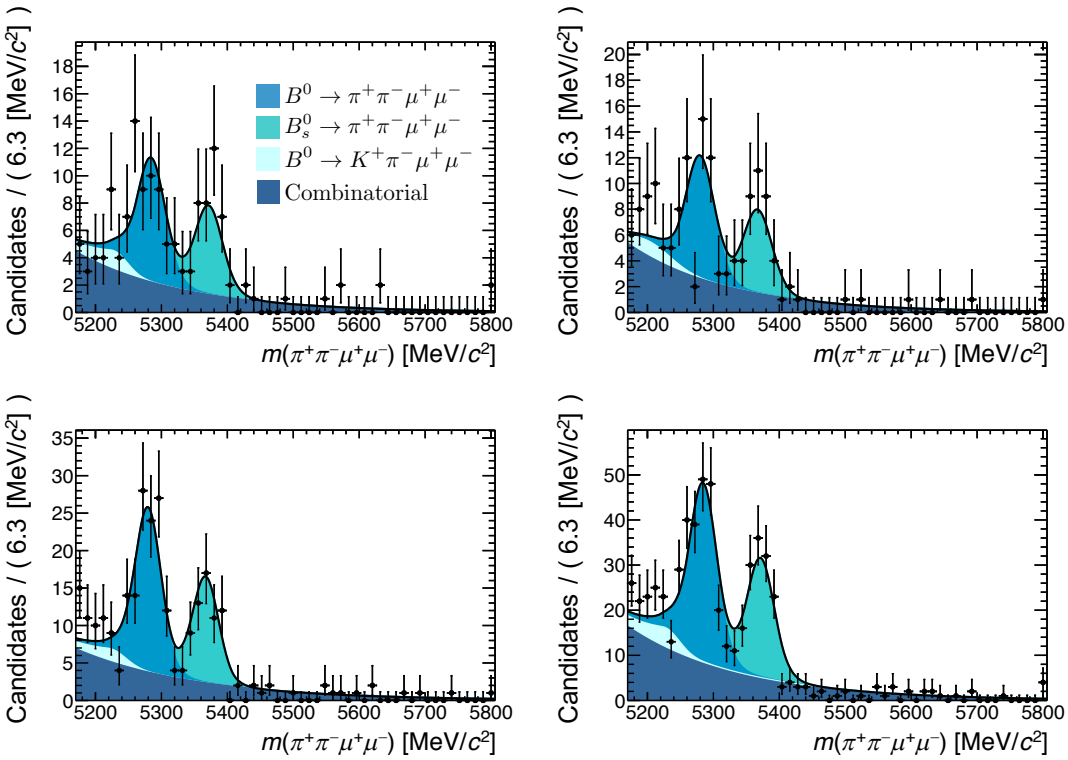

Figure 8.19: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates within the full q^2 region from the data. The result of the simultaneous fit is compared to the data. The top left plot shows the fit to the 2011+2012 dataset, the top right the fit to the 2015+2016 dataset and the bottom left the fit to the 2017+2018 dataset. The bottom right plot combines the datasets and the fits for visualisation.

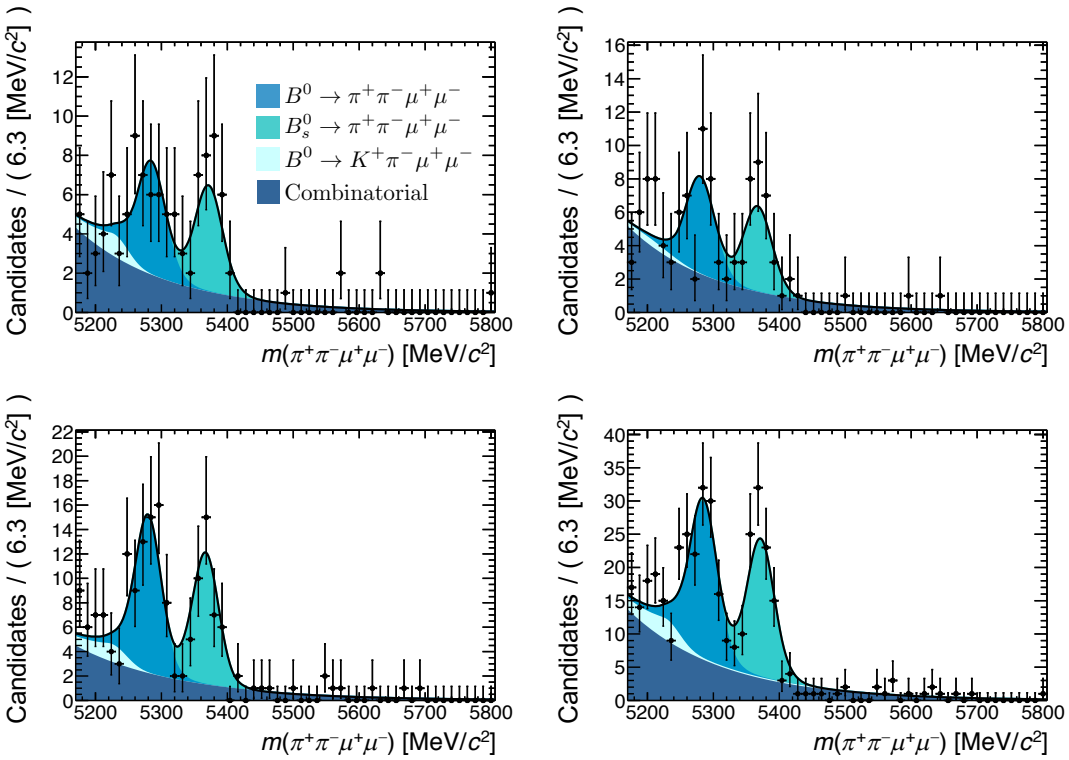

Figure 8.20: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates within the low q^2 region from the data. The result of the simultaneous fit is compared to the data. The top left plot shows the fit to the 2011+2012 dataset, the top right the fit to the 2015+2016 dataset and the bottom left the fit to the 2017+2018 dataset. The bottom right plot combines the datasets and the fits for visualisation.

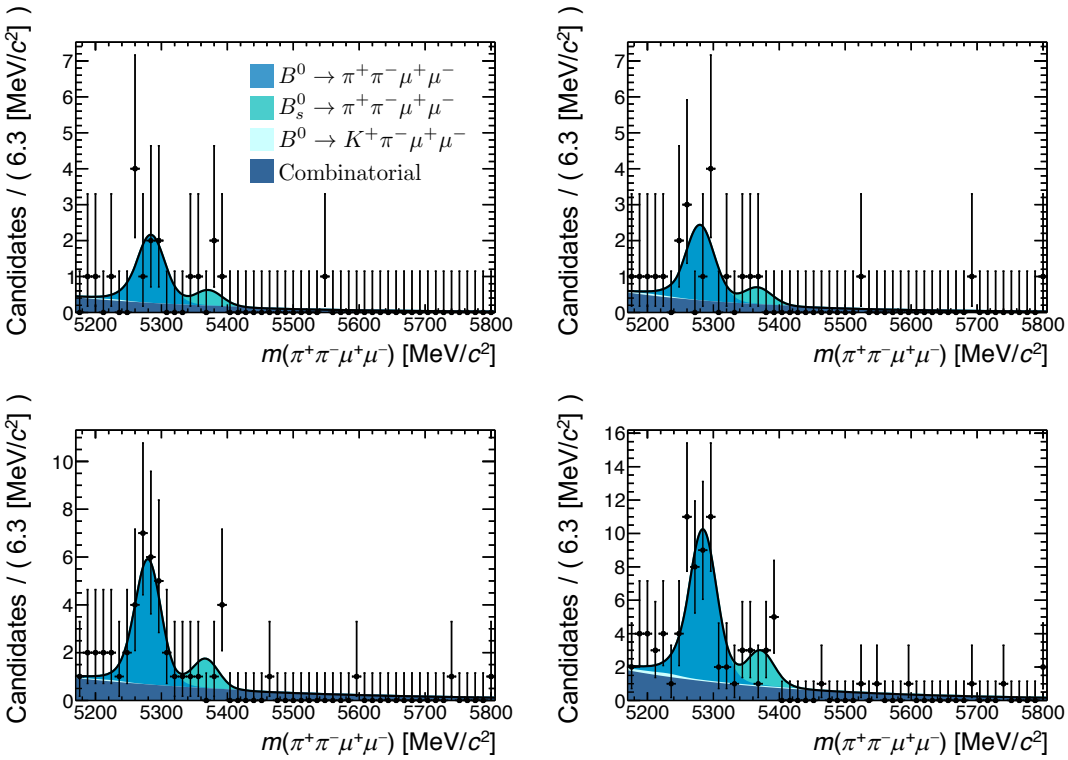

Figure 8.21: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates within the high q^2 region from the data. The result of the simultaneous fit is compared to the data. The top left plot shows the fit to the 2011+2012 dataset, the top right the fit to the 2015+2016 dataset and the bottom left the fit to the 2017+2018 dataset. The bottom right plot combines the datasets and the fits for visualisation.

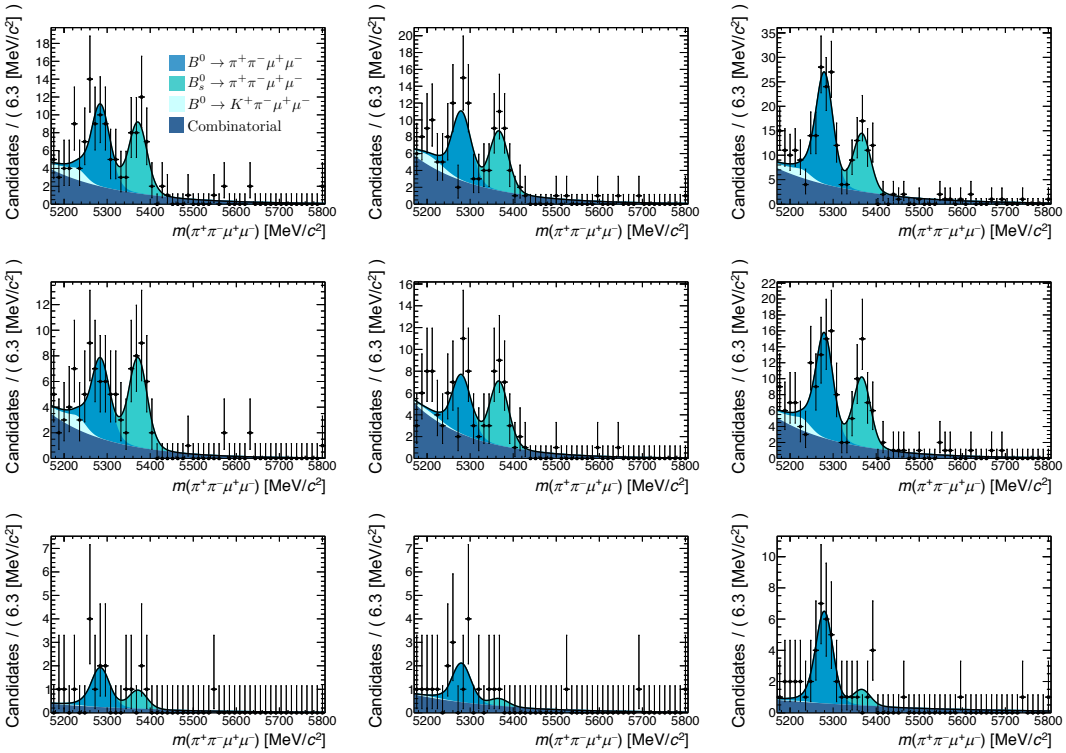

Figure 8.22: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates from the data compared to the invariant mass fit when fitting each dataset separately. The left column shows the fits to the 2011+2012 dataset, the middle column the fits to the 2015+2016 dataset and the right column the fit to the 2017+2018 dataset. The top row shows the results for the full q^2 region, the middle row the results for the low q^2 region and the bottom row the results for the high q^2 region.

Table 8.5: Parameters from the simultaneous fit to the 2011+2012, 2015+2016 and 2017+2018 datasets for the full, low and high q^2 ranges. The rare mode yields in blue are not varied within the fit and are calculated from the branching fraction and event sensitivity (κ) parameters. These yields are included here for convenient comparison.

Dataset	Parameter	Full q^2	Low q^2	High q^2
	$B(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	2.34 ± 0.26	1.02 ± 0.14	0.37 ± 0.07
	$B(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	5.33 ± 0.64	3.18 ± 0.41	0.28 ± 0.13
2011+2012	$\kappa(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.057 ± 0.004	0.041 ± 0.003	0.041 ± 0.003
	$N(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	41 ± 5	25 ± 4	9 ± 2
	$\kappa(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.176 ± 0.016	0.129 ± 0.012	0.132 ± 0.012
	$N(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	30 ± 5	25 ± 4	2 ± 1
	λ	-0.006 ± 0.001	-0.007 ± 0.001	-0.004 ± 0.002
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	8 ± 1	6 ± 1	0 ± 1
	N_{bkg}	57 ± 11	46 ± 10	7 ± 4
2015+2016	$\kappa(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.052 ± 0.003	0.037 ± 0.003	0.033 ± 0.002
	$N(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	45 ± 6	28 ± 4	11 ± 2
	$\kappa(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.166 ± 0.014	0.127 ± 0.011	0.108 ± 0.010
	$N(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	32 ± 5	25 ± 4	3 ± 1
	λ	-0.007 ± 0.001	-0.007 ± 0.001	-0.005 ± 0.002
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	7 ± 1	2 ± 1	0 ± 1
	N_{bkg}	68 ± 11	56 ± 10	10 ± 4
2017+2018	$\kappa(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.024 ± 0.002	0.017 ± 0.001	0.017 ± 0.002
	$N(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	97 ± 14	60 ± 9	22 ± 5
	$\kappa(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.082 ± 0.006	0.061 ± 0.005	0.053 ± 0.005
	$N(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$	65 ± 9	52 ± 8	5 ± 3
	λ	-0.006 ± 0.001	-0.006 ± 0.001	-0.003 ± 0.001
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	10 ± 1	9 ± 1	0 ± 1
	N_{bkg}	105 ± 14	60 ± 11	22 ± 6

Table 8.6: Parameters from the separate fits to the 2011+2012, 2015+2016 and 2017+2018 datasets for the full, high and low q^2 ranges. The rare mode yields in blue are not varied within the fit and are calculated from the branching fraction and event sensitivity (κ) parameters. These yields are included here for convenient comparison.

Dataset	Parameter	Full q^2	Low q^2	High q^2
2011+2012	$B(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	2.51 ± 0.56	1.20 ± 0.33	0.33 ± 0.15
	$B(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	7.26 ± 1.53	4.69 ± 1.02	0.53 ± 0.32
	$\kappa(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.058 ± 0.004	0.041 ± 0.003	0.041 ± 0.003
	$N(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	43 ± 10	29 ± 8	8 ± 4
	$\kappa(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	0.186 ± 0.017	0.137 ± 0.012	0.133 ± 0.012
	$N(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	39 ± 9	34 ± 8	4 ± 2
	λ	-0.006 ± 0.001	-0.007 ± 0.001	-0.004 ± 0.002
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	8 ± 1	6 ± 1	0 ± 1
	N_{bkg}	52 ± 11	40 ± 10	7 ± 4
2015+2016	$B(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	1.98 ± 0.49	0.96 ± 0.30	0.27 ± 0.13
	$B(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	5.94 ± 1.34	3.80 ± 0.92	0.18 ± 0.22
	$\kappa(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.051 ± 0.003	0.037 ± 0.003	0.032 ± 0.002
	$N(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	39 ± 10	26 ± 8	6 ± 7
	$\kappa(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	0.169 ± 0.014	0.129 ± 0.011	0.108 ± 0.010
	$N(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	35 ± 8	29 ± 8	2 ± 2
	λ	-0.007 ± 0.001	-0.008 ± 0.001	-0.005 ± 0.001
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	7 ± 1	2 ± 1	0 ± 1
	N_{bkg}	70 ± 13	55 ± 11	12 ± 5
2017+2018	$B(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	2.44 ± 0.35	0.99 ± 0.18	0.45 ± 0.12
	$B(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	4.17 ± 0.79	2.29 ± 0.47	0.25 ± 0.17
	$\kappa(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	0.024 ± 0.002	0.017 ± 0.001	0.017 ± 0.002
	$N(B^0 \to \pi^+\pi^-\mu^+\mu^-)$	102 ± 17	58 ± 11	26 ± 8
	$\kappa(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	0.077 ± 0.007	0.058 ± 0.005	0.053 ± 0.005
	$N(B^0_s \to \pi^+\pi^-\mu^+\mu^-)$	54 ± 11	39 ± 9	5 ± 3
	λ	-0.006 ± 0.001	-0.006 ± 0.001	-0.003 ± 0.001
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	10 ± 1	9 ± 1	0 ± 1
	N_{bkg}	108 ± 15	64 ± 12	20 ± 7

CHAPTER 9

Systematic Uncertainties

> *If you thought that science was certain - well, that is just at error on your part.*
>
> Richard Feynman,

This chapter discusses the various sources of systematic uncertainty which affect the precision of the final branching fraction result. The motivation for performing the branching fraction measurements relative to a control mode was that many of the major sources of systematic uncertainty cancel in the ratio. However, some sources of systematic uncertainty are not expected to fully cancel in this way and so these must be included in the analysis.

This chapter explains the way in which the different sources of systematic uncertainty are estimated in the analysis. A total systematic uncertainty on the branching fraction result is then determined by summing the separate uncertainties in quadrature. Depending on the method of estimation, the uncertainties are calculated either for each of the three datasets separately, or, using the three datasets simultaneously. When calculating the uncertainties for each separate dataset, the total systematic uncertainty is determined by summing the individual uncertainties together using a weighted average using the integrated luminosity information.

Ultimately, the rarity of the studied processes means that the dominant source of uncertainty to the analysis is statistical and the uncertainty on the final results would only become systematically dominated with the inclusion of future LHCb data.

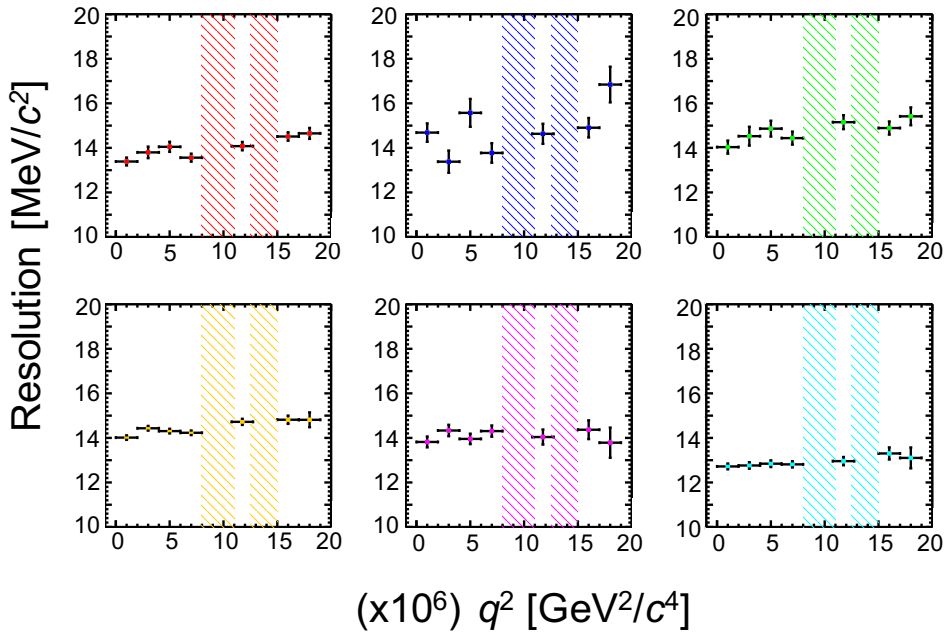

Figure 9.1: Resolutions from the double CB + Gaussian fit to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ simulation in bins of q^2. The top row shows the fits, from left to right, for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ simulation for the 2011+2012, 2015+2016 and 2017+2018 samples. The bottom row shows the equivalent plots for the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ simulation.

9.1 Signal Line Shape Variation

The first systematic uncertainty considered arises from the line shape which is used to describe the $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ signal peak within the final rare mode invariant mass fits. The largest uncertainty here is due to the variation of the resolution of the signal peak as a function of q^2 and the fact that the q^2 distribution of both rare modes is not well known experimentally.

To determine the variation of the resolution with q^2, invariant mass fits are performed on the $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$ simulation samples in the different bins of q^2. Then, the width of the CB shapes which describe the signal peaks are taken as a proxy for the total resolution. The variation of the resolution with q^2, for the three datasets can be seen in Figure 9.1 where a dependence of the resolution on q^2 can be seen. This trend is due to the increasing importance of the muon momentum resolution towards the total resolution at higher invariant masses.

To convert this dependence into a systematic uncertainty a pseudo-experiment study is performed where the width of the signal peaks within the fit is modified and then the full fitting model is refit to a generated dataset. The generated dataset is created using the original,

non-modified signal width. Two sets of pseudo-experiments are then performed. The first set increases the model width by $0.5\,\mathrm{MeV}/c$, and the second set reduces the signal width by the same amount. The resulting systematic uncertainty is then the difference between the mean of the branching fraction distribution where the signal width was increased and the case where it was decreased, as a percentage of the expected branching fractions. The modification of $\pm 0.5\,\mathrm{MeV}/c$ is chosen as a reasonable estimate of the average change in the resolution which is seen across the q^2 range from Figure 9.1. The overall size of the systematic uncertainty is evaluated to be 0.2% for the B^0 mode and 0.15% for the B_s^0 mode showing that the size of the effect is very small.

9.2 Inclusion of $B_s^0 \to \eta' \mu^+ \mu^-$

The partially reconstructed background $B_s^0 \to \eta' \mu^+ \mu^-$ was identified within Chapter 5 as a potential source of peaking background which was potentially relevant for the analysis. However, the expected yield of this background was judged to be small enough to neglect this component from the final invariant mass fits. Here, the systematic effect of including the background in the final mass fits is considered and is estimated using a pseudo-experiment study.

To perform the study 1000 pseudo-experiments are generated. For each, the generated dataset is created using the final rare mode PDF and the corresponding expected yields. The η' component is then added to the pseudo-experiment dataset by generating new candidates using the shape of the $B_s^0 \to J/\psi \eta'$ component taken from the resonant mode fits. The number of generated candidates is taken from the expected $B^0 \to \eta' \mu^+ \mu^-$ yield. The total generated dataset is then fit with the model which does not include the η' process component. After running the pseudo-experiments the difference between the mean of the resulting branching fraction distribution and the expected branching fraction provides an estimation of the systematic uncertainty. The systematic uncertainties are estimated to be 1.46%, 2.98% and 2.32% for the 2011+2012, 2015+2016 and 2017+2018 datasets respectively for the B^0 mode and 1.05%, 0.93% and 0.08% respectively for the B_s^0 mode.

9.3 Efficiency Variations

The statistical uncertainties on the different efficiency values which are used within the analysis were calculated using a bootstrapping technique as was discussed in Chapter 7. The total selection efficiency uncertainty, in addition to the geometrical acceptance efficiency and PID efficiency uncertainty, are added in quadrature to determine the total uncertainty. A systematic uncertainty is assigned to account for the size of the effect of the variation of this total efficiency on the final branching fraction results.

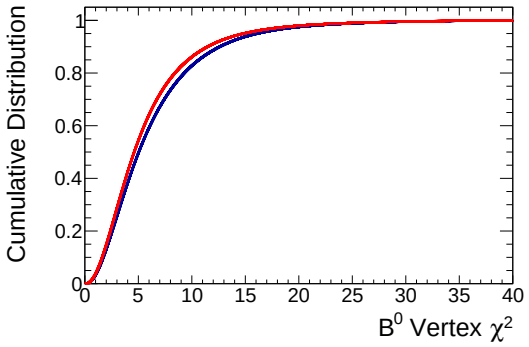

Figure 9.2: Cumulative distribution function for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC sample in red and the *sWeighted* control mode sample in blue. The comparison is made with the 2017+2018 sample with similar plots seen for the other two datasets.

Using Equation 4.2 from Chapter 4 the efficiency ratio term is fluctuated within its uncertainty while the yields, fragmentation fractions and control mode branching fractions within the equation are held constant. This is repeated 1000 times to re-evaluate the branching fractions for each fluctuation. The standard deviation of the resulting branching fraction distribution is then taken as the systematic uncertainty.

The size of the effect of the efficiency variation on the final results are found to be very small. For the B^0 mode the resultant systematic uncertainty is estimated to be 0.29% for the 2011+2012 dataset, 0.68% for the 2015+2016 dataset and 0.61% for the 2017+2018 dataset. For the B_s^0 mode the equivalent systematics are evaluated to be $0.36\%, 0.81\%$ and 0.91% for the 2011+2012, 2015+2016 and 2017+2018 datasets respectively.

9.4 Vertex χ^2 Re-sampling

Within the analysis differences in the B^0 vertex fit χ^2 distribution between data and MC are addressed using the kinematic weighting procedure from Chapter 5. An alternative method to deal with this discrepancy is investigated which involves re-sampling the B^0 vertex fit χ^2 variable using a data-driven method. This method compares the normalised cumulative distribution function (CDF) of the B^0 vertex fit χ^2 variable between data and MC. The data distribution is extracted by applying *sWeights* to the control mode data sample. Then, for each candidate in the MC sample, the value of the B^0 vertex fit χ^2 variable is used to look up the corresponding CDF which is mapped to the CDF for the data to obtain a new value to be used in the MC sample. Examples of the CDF distributions for data and MC can be seen in Figure 9.2 while the results of the re-sampling can be seen in Figure 9.3.

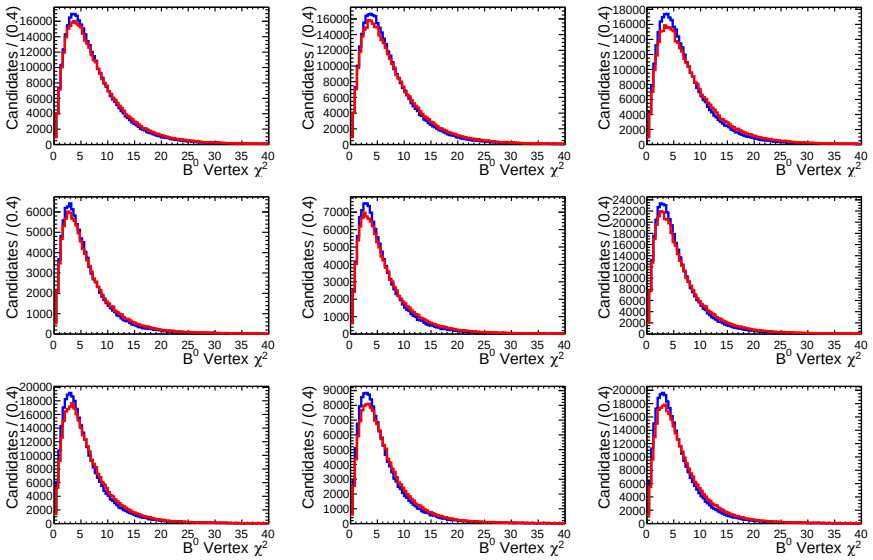

Figure 9.3: Results of the re-sampling procedure for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC (left column), $B^0_s \to \pi^+\pi^-\mu^+\mu^-$ (middle column) and $B^0 \to J/\psi\, K^{*0}$ (right column) for 2011+2012 (top row), 2015+2016 (middle row) and 2017+2018 (bottom row). The original distributions are shown in blue and the re-sampled distributions in red.

This alternative method of correction mainly affects the performance of the BDT within the analysis as the BDT uses this variable to help in the classification of signal and background candidates. Therefore, the BDT efficiency is re-evaluated for the rare mode samples and the control mode samples with both the re-sampling and BDT classifier re-applied. The difference in the calculated BDT efficiency between the two methods is then propagated through Equation 4.2 to estimate a systematic uncertainty on the final branching fraction results. The size of the effect is small with uncertainties of 0.12%, 0.25%, 1.62% estimated for the 2011+2012, 2015+2016 and 2017+2018 datasets for the B^0 mode. The equivalent uncertainties for the B^0_s mode are also small and are found to be 0.49%, 0.20%, 0.08%.

9.5 Decay Model Variations

The decay model weighting from Chapter 6 is a potential source of systematic uncertainties within the analysis and several different sources are considered. These include the values of the Wilson coefficients used within the model, the value of the form factors used within the model and the inclusion of different resonances within the model. In order to estimate these uncertainties, the decay model is run twice, once with the default model and once with an alteration. The full

set of efficiencies is then estimated for both these scenarios with the difference in the efficiencies propagated through Equation 4.2 to extract an estimated uncertainty on the rare mode branching fractions.

The Wilson coefficient uncertainties are evaluated by considering two alternative sets of coefficient values. The first varies the value of C_9, setting C_9 to 3.211 instead of the default value of 4.211. The second varies both the value of C_9 and C_{10}, setting C_9 to 3.511 and C_{10} to -3.403 instead of -4.103. These variations are driven by global fits which show that $C_{9\text{NP}} = -1.0$ and $C_{9\text{NP}} = -C_{10\text{NP}} = -0.7$ [116] is best able to explain the measurements of $b \to s\mu^+\mu^-$ observables. These two forms of Wilson coefficient variation are not independent of one another. Therefore, to prevent the double counting of systematic uncertainties, it is the C_9 variation which is ultimately chosen as the final systematic as this normally results in the larger uncertainty.

The uncertainty from the form factors is evaluated by fluctuating the form factors within their uncertainties. These uncertainties are taken from the relevant papers in Chapter 6.

The final uncertainty is evaluated by applying the decay model including a single resonance contribution, instead of multiple resonance contributions. This single resonance is taken to be the $\rho(770)$ for the B^0 mode and the $f_0(980)$ for the B_s^0 mode. The inclusion of a single resonance mainly affects the mass and q^2 efficiencies within the analysis by altering the shape of the $m(\pi^+\pi^-)$ and q^2 distributions and this is expected to result in a relatively large systematic uncertainty value compared to some of the other variations made.

The systematic uncertainties relevant to the decay model can be found in the tables at the end of this chapter. All other sources of systematic uncertainty evaluated within this chapter are evaluated only for the full q^2 regime and the assumption is made that the size of the uncertainty should be generally invariant across the different q^2 regions of interest. This assumption is not made for the decay model uncertainties however. This is because the decay model mainly effects the size of the calculated mass and q^2 efficiencies and these values vary between the different q^2 regimes. Therefore, the decay model systematic uncertainties are recalculated for both the high and low q^2 regimes and these values are also found at the end of this chapter.

9.6 Kinematic Weighting

The kinematic weighting from Chapter 5 is evaluated as a potential source of systematic uncertainty in two ways. Firstly, a discrete check is made to test the impact of using a particular binning scheme when weighting the three variables. Within the analysis the p_{T} distribution weighting is performed with 51 bins whilst the vertex χ^2 weighting uses 50 bins and the occupancy uses 44 bins for nLongTracks and 50 bins for nSPDHits. These bins are of varying sizes to ensure that each bin has a large enough number of candidates and that there is a smooth variation between

adjacent bins. An alternative binning scheme is used to evaluate the systematic which uses 20 bins for each weighted variable where each bin is of the same size. Then, the weighting is repeated and the efficiency values, and subsequent branching fraction estimates, recalculated to determine the systematic. The systematic effect is found to be small with the B^0 mode showing a 0.15%, 1.69% and 0.91% effect, and the B_s^0 mode showing a 0.41%, 1.18% and 0.91% effect for the 11+12, 15+16 and 17+18 samples respectively.

A second check is made to account for any correlations between the three kinematic variables. This check is made using a bootstrapping technique whereby the calculation of the individual weights for the three variables is repeated a large number of times where different variations of the data and MC distributions are used to calculate the weights with each bootstrap. Here, the bootstrapping is performed by giving each candidate within the data and MC samples a weight generated from a Poisson distribution with a mean of 1. The analysis efficiencies are re-evaluated for each bootstrap sample and the width of the resulting branching fraction distribution is taken as the uncertainty. The size of this systematic effect is very small and is evaluated to be 0.04%, 0.09% and 0.09% for the B^0 mode and 0.03%, 0.09% and 0.06% effect for the B_s^0 modes for the 11+12, 15+16 and 17+18 datasets respectively.

9.7 B_s^0 Lifetime Weighting

Due to the fast nature of B_s^0 meson oscillations and the non-zero width difference between the light and heavy mass eigenstates, the B_s^0 lifetime seen within the data lies between the lifetimes of the two mass eigenstates. This could present a source of uncertainty for the B_s^0 branching fraction measurement. To estimate the size of this uncertainty each candidate is weighted to rescale the B_s^0 lifetime distribution to either the light or heavy mass eigenstate lifetime distribution. The efficiencies are then recalculated for the weighted datasets and the new efficiency is propagated through to obtain a difference in the predicted branching fraction measurement. This difference is then taken as the systematic uncertainty with the resulting values found in the tables at the end of the chapter.

9.8 Systematic Uncertainties Summary

Tables showing the systematic uncertainties for the B^0 and B_s^0 rare modes as a percentage of the branching fraction results are given in Table 9.1. These uncertainties are estimated for the full q^2 range. The systematic uncertainties for the low and high q^2 range are assumed to be the same apart from the decay model systematic uncertainties. These are reevaluated for the low and high q^2 regimes and these values are shown in Table 9.2 and Table 9.3 respectively.

The largest source of systematic uncertainty on the full $q^2 B^0$ branching fraction comes from the decay model when weighting with a single resonance. This is understandable as the decay model weighting changes the fraction of candidates falling into the J/ψ and $\psi(2S)$ mass windows. Within both the low and high q^2 regions the size of this uncertainty is seen to be much smaller.

The largest source of systematic uncertainty on the B_s^0 branching fraction originates from the weighting of the lifetime to match either the heavy or light eigenstate lifetimes. The lifetime information is used to train the BDT classifier during the selection and it is this aspect of the analysis which drives the variation.

The largest systematic uncertainty calculated, at 7.7%, for the full q^2 range of the B^0 decay, is still much smaller than the statistical uncertainty on the final branching fraction results which are approximately 12% for both B^0 and B_s^0 mode. The size of the asymmetry between the systematic and statistical uncertainties results in almost complete statistical uncertainty domination when adding the two sources in quadrature. Therefore, if trying to repeat this analysis with the aim of improving the precision of the final result, the first priority would be to increase the size of the datasets used. This is achievable in the future with new LHCb data-taking runs.

Table 9.1: Systematic uncertainties for the rare modes estimated for the full q^2 range. The total systematic uncertainty is obtained by summing the separate uncertainties in quadrature. The uncertainties for the different datasets are combined using integrated luminosity information.

Decay	Systematic Error	11+12	15+16	17+18	11-18
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Signal Resolution	–	–	–	0.20%
	η' background	1.46%	2.98%	2.32%	–
	Efficiency Variations	0.29%	0.68%	0.61%	–
	Vertex χ^2 weighting	0.12%	0.25%	1.62%	–
	Alternate C_9	0.79%	2.27%	2.44%	–
	Alternate C_9, C_{10}	0.54%	2.10%	2.30%	–
	Single Resonance	7.22%	5.36%	6.28%	–
	Form Factor Variation	3.13%	1.42%	3.33%	–
	Kinematic Weighting: Binning Scheme	0.15%	1.69%	0.91%	–
	Kinematic Weighting: Bootstrap	0.04%	0.09%	0.08%	–
	Total Systematic:	**7.66%**			
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	Signal Resolution	–	–	–	0.153%
	Heavy Eigenstate weight	2.30%	2.54%	2.12%	–
	Light Eigenstate weight	2.84%	2.52%	2.54%	–
	η' background	1.05%	0.93%	0.078%	–
	Efficiency Variations	0.36%	0.80%	0.91%	–
	Vertex χ^2 weighting	0.49%	0.20%	0.08%	–
	Alternate C_9	0.83%	1.37%	0.84%	–
	Alternate C_9,C_{10}	0.84%	1.26%	0.79%	–
	Single Resonance	2.74%	1.36%	1.54%	–
	Form Factor Variation	0.85%	1.49%	1.45%	–
	Kinematic Weighting: Binning Scheme	0.41%	1.18%	0.91%	–
	Kinematic Weighting: Bootstrap	0.03%	0.09%	0.06%	–
	Total Systematic:	**4.46%**			

Table 9.2: Decay model systematic uncertainties estimated for low q^2. All other sources of systematic uncertainties are assumed to be the same as for the full q^2 range and the total systematic uncertainty includes these values.

Decay	Systematic Error	11+12	15+16	17+18
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Alternate C_9	2.33%	2.39%	2.21%
	Alternate C_9, C_{10}	1.96%	1.88%	1.93%
	Form Factor Variation	1.89%	1.85%	0.97%
	Single Resonance	1.47%	0.27%	0.99%
	Total Systematic:		**3.84%**	
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	Alternate C_9	1.20%	1.15%	1.22%
	Alternate C_9, C_{10}	1.23%	1.12%	1.20%
	Form Factor Variation	0.28%	0.38%	0.21%
	Single Resonance	1.74%	0.70%	1.09%
	Total Systematic:		**4.71%**	

Table 9.3: Decay model systematic uncertainties estimated for high q^2. All other sources of systematic uncertainties are assumed to be the same as for the full q^2 range and the total systematic uncertainty includes these values.

Decay	Systematic Error	11+12	15+16	17+18
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	Alternate C_9	0.10%	0.03%	0.16%
	Alternate C_9, C_{10}	0.10%	0.23%	0.08%
	Form Factor Variation	2.31%	2.23%	3.31%
	Single Resonance	0.61%	0.54%	1.68%
	Total Systematic:		**3.85%**	
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	Alternate C_9	1.02%	1.08%	0.85%
	Alternate C_9, C_{10}	1.04%	1.01%	0.87%
	Form Factor Variation	1.96%	1.76%	3.06%
	Single Resonance	0.12%	0.20%	2.23%
	Total Systematic:		**4.63%**	

CHAPTER 10

Results and Conclusions

“However beautiful the strategy, you should occasionally look at the results”

Winston Churchill,

10.1 Invariant Mass Fit Results

The final results of the rare mode invariant mass fits can be found in Table 10.1 where the uncertainties are statistical and the significances are calculated using Wilk's theorem [171]. The results of the fits are presented for all three of the q^2 regions and for both the simultaneous fits and the individual fits to the three datasets. The likelihood profiles from the final invariant mass fits can be found in Figure 10.1. The changes in likelihood are seen to show a parabolic behaviour which have minima corresponding to the branching fraction values from Table 10.1, indicating that the values converged upon are accurate.

The branching fraction results are visualised in Figure 10.2, Figure 10.3 and Figure 10.4. For the full q^2 range, the branching fraction results from the previous LHCb analysis are included for comparison. As can be seen from the figure, a good level of agreement is found between the individual fits to the different datasets. This is seen to be true for all of the q^2 regions. When comparing to the previous LHCb values, the most suitable sets of results for comparison are the Run 1 results from this thesis across the full q^2 range. As the candidate selection is different between this analysis and the previous analysis, the uncertainties are not completely correlated and it can be seen that both the B^0 and B_s^0 branching fraction values are seen to agree with the

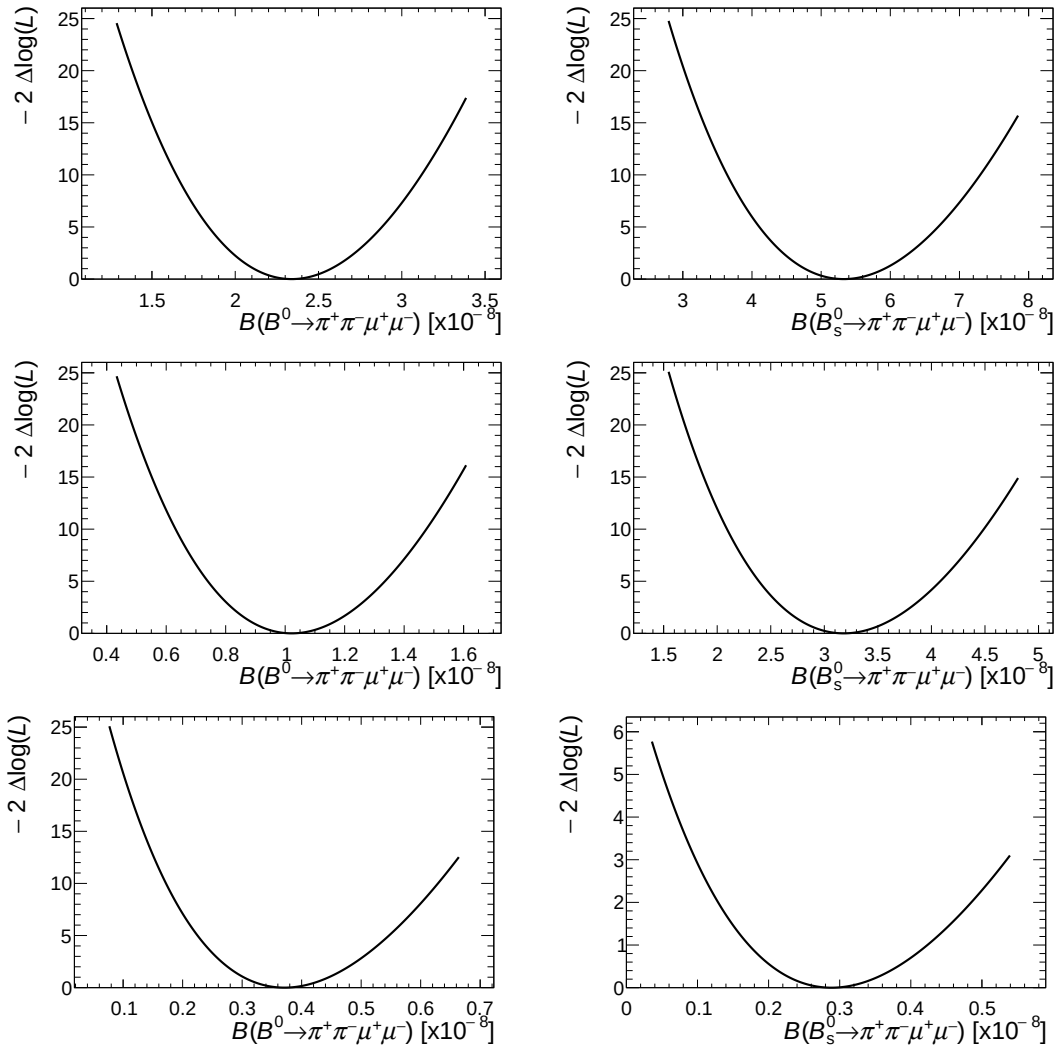

Figure 10.1: The change in log-likelihood, $-2\Delta\log L$, from the best fit point from the simultaneous fits to the three datasets for the full q^2 range (top), the low q^2 range (middle) and the high q^2 range (bottom) when varying the B^0 branching fraction (left column) and the B_s^0 branching fraction (right column).

q^2 range	Dataset	$B(B^0 \to \pi^+\pi^-\mu^+\mu^-)$ $[10^{-8}]$	B^0 significance $[\sigma]$	$B(B_s^0 \to \pi^+\pi^-\mu^+\mu^-)$ $[10^{-8}]$	B_s^0 significance $[\sigma]$
$[0.1 - 19.0]\,\mathrm{GeV}^2/c^4$	Full	2.34 ± 0.26	12.5	5.33 ± 0.64	12.6
	11+12	2.51 ± 0.56	5.7	7.26 ± 1.53	7.4
	15+16	1.98 ± 0.49	5.0	5.94 ± 1.34	6.4
	17+18	2.44 ± 0.35	10.0	4.17 ± 0.79	8.1
$[0.1 - 8.0]\,\mathrm{GeV}^2/c^4$	Full	1.02 ± 0.14	9.5	3.18 ± 0.41	11.8
	11+12	1.20 ± 0.33	4.4	4.69 ± 1.02	7.3
	15+16	0.96 ± 0.30	3.9	3.80 ± 0.92	6.0
	17+18	0.99 ± 0.18	7.4	2.29 ± 0.47	7.5
$[15.0 - 19.0]\,\mathrm{GeV}^2/c^4$	Full	0.37 ± 0.07	6.7	0.29 ± 0.13	2.8
	11+12	0.33 ± 0.15	3.0	0.53 ± 0.32	2.2
	15+16	0.27 ± 0.13	2.4	0.18 ± 0.22	1.0
	17+18	0.45 ± 0.12	5.5	0.25 ± 0.17	1.8

Table 10.1: Branching fraction results for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay and the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decay mode evaluated for the different datasets, the combined dataset and the different q^2 regions. The statistical significance of each measurement is calculated using Wilk's theorem. The quoted uncertainties are statistical.

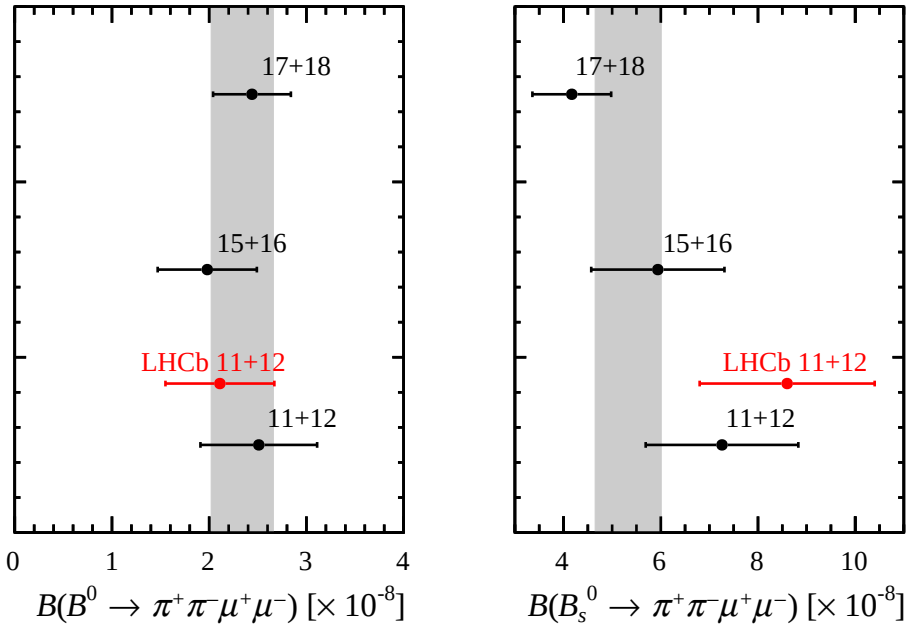

Figure 10.2: Branching fraction results from the rare mode invariant mass fits for the full q^2 region, separated by data-taking period. The results of the fits to the individual datasets are shown as data points while the result of the simultaneous fit is shown as a grey band. The uncertainties are the result of adding the statistical and systematic uncertainties in quadrature. The red data points correspond to the previous LHCb measurements.

previous LHCb values to within 1 σ. Comparing the previous LHCb results to the full Run 1 + Run 2 results from this analysis shows a 0.3 σ tension between the two B^0 mode results and a 1.6 σ tension between the two B_s^0 mode results.

The low and high q^2 information can be combined into a single plot showing the binned branching fraction results as a function of q^2. This can be seen in Figure 10.5, where the branching fraction for the B_s^0 mode is shown in red and the equivalent for the B^0 mode in black. The relative drop in the branching fraction from low to high q^2 is seen to be far larger for the B_s^0 mode than for the B^0 mode. This is due to the main resonance through which the $\pi^+\pi^-$ system originates. The B_s^0 decay primarily proceeds via the $f_0(980)$ resonance which is significantly heavier than the $\rho^0(770)$ resonance. As a result, when measuring the branching fraction at high q^2, the possible phase space is smaller for the B_s^0 mode than for the B^0 mode. In addition there is an enhancement of vector form-factors at high values of q^2 which is not present within the equivalent scalar form-factors.

Referring back to Table 10.1, the statistical significance of the different fits show that the

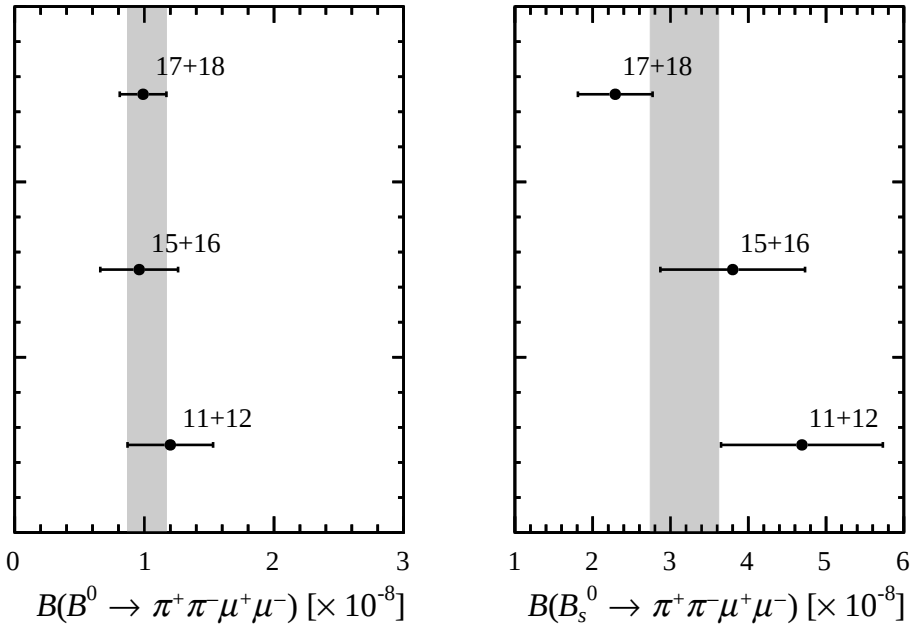

Figure 10.3: Branching fraction results from the rare mode invariant mass fits for the low q^2 region of $q^2 \in [0.1 - 8.0]\,\mathrm{GeV}^2/c^4$. The results of the fits to the individual datasets are shown as data points while the result of the simultaneous fit is shown as a grey band. The uncertainties are the result of adding the statistical and systematic uncertainties in quadrature.

existence of the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ mode is confirmed with a statistical significance of 12.5 σ in the full q^2 range, a significance of 9.5 σ in the low q^2 range and a significance of 6.7 σ in the high q^2 range. These are therefore the first confirmed observations for this decay mode in all three q^2 regions. The $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ mode has already been confirmed to exist within the full q^2 region by the previous LHCb result. However the level of statistical significance has now increased from 7.2 σ to 12.6 σ. In addition, this analysis provides a first measurement of the B_s^0 mode in the low q^2 range, with a significance of 11.8 σ. The significance of the B_s^0 mode branching fraction within the high q^2 bin is 2.8 σ, falling short of the 5 σ criterion for evidence.

The pseudo-experiment study from Chapter 8 showed that a very small level of bias was detected in the fitting procedure prior to the completion of the final rare mode invariant mass fits. This bias, albeit small, was consistent and so the final branching results should account for this small effect. When applying this shift and also including the systematic uncertainties to the branching fractions, the final results, as seen in Table 10.2 are obtained.

By assuming that the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ signal is dominated by $B^0 \to \rho^0(770)\mu^+\mu^-$ and that $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ is dominated by $B_s^0 \to f_0(980)\mu^+\mu^-$ decays a brief, qualitative

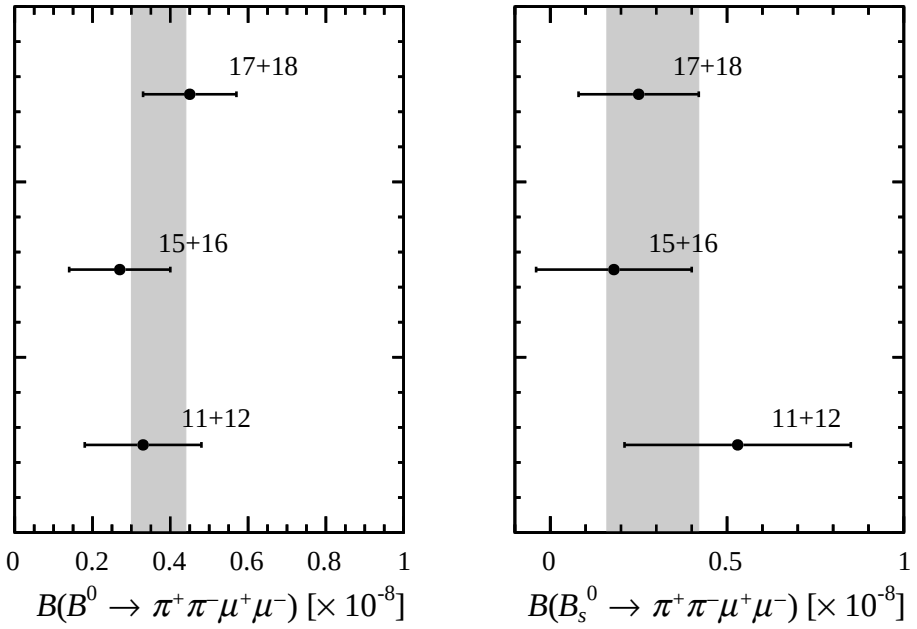

Figure 10.4: Branching fraction results from the rare mode invariant mass fits for the low q^2 region of $q^2 \in [15.0 - 19.0]\,\mathrm{GeV}^2/c^4$. The results of the fits to the individual datasets are shown as data points while the result of the simultaneous fit is shown as a grey band. The uncertainties are the result of adding the statistical and systematic uncertainties in quadrature.

discussion can be had comparing the branching fraction results to theoretical predictions. The branching fraction result across the full q^2 range for the B^0 mode is seen to agree well with the SM predictions taken from [84] which predicts $B(B^0 \to \rho^0\mu^+\mu^-) \sim 2.4 \times 10^{-8}$ with a 30% uncertainty. The B^0 mode branching fraction result for the low q^2 region is seen to also broadly agree with the SM predictions taken from [86] where the prediction is given for the narrower q^2 range $q^2 \in [0.0 - 6.0]\,\mathrm{GeV}^2/c^4$.

The branching fraction result for the B_s^0 mode is seen to disfavour the SM prediction taken for the full q^2 range from [123] but shows no significant tension with the leading order SM prediction from [188]. The B_s^0 mode branching fraction also shows broad agreement with the SM predictions for the full, low and high q^2 regions taken from [125]. A more complete comparison with theoretical predictions requires the identification of which states contribute towards the $\pi^+\pi^-$ system, which is further discussed in Section 10.2. Furthermore, the precision of the theoretical predictions are dominated by large hadronic form-factor uncertainties. A future comparison of the rare modes to theory would benefit from a reduction within these uncertainties in addition to additional experimental data to reduce the statistical uncertainty which dominates

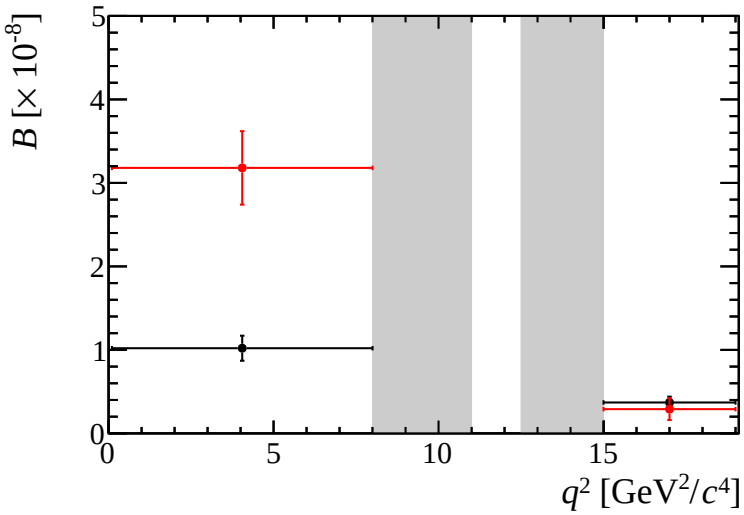

Figure 10.5: Branching fraction results for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay (black) and the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decay (red) mode in bins of q^2. The uncertainties are calculated by adding the statistical and systematic uncertainties in quadrature.

Decay	q^2 range	Branching Fraction [10^{-8}]	Significance
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	$q^2 \in [0.1 - 19.0]\,\text{GeV}^2/c^4$	2.35 ± 0.26 (stat) ± 0.18 (syst)	12.5 σ
	$q^2 \in [0.1 - 8.0]\,\text{GeV}^2/c^4$	1.03 ± 0.14 (stat) ± 0.04 (syst)	9.5 σ
	$q^2 \in [15.0 - 19.0]\,\text{GeV}^2/c^4$	0.37 ± 0.07 (stat) ± 0.01 (syst)	6.7 σ
$B_s^0 \to \pi^+\pi^-\mu^+\mu^-$	$q^2 \in [0.1 - 19.0]\,\text{GeV}^2/c^4$	5.35 ± 0.64 (stat) ± 0.24 (syst)	12.6 σ
	$q^2 \in [0.1 - 8.0]\,\text{GeV}^2/c^4$	3.19 ± 0.41 (stat) ± 0.15 (syst)	11.8 σ
	$q^2 \in [15.0 - 19.0]\,\text{GeV}^2/c^4$	0.30 ± 0.13 (stat) ± 0.01 (syst)	2.8 σ

Table 10.2: Final branching fraction results for $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$. The significance is calculated using Wilk's theorem.

the uncertainty on the experimental results.

10.2 Evidence for $B^0 \to \rho^0(770)\mu^+\mu^-$ and $B_s^0 \to f_0(980)\mu^+\mu^-$

As was discussed in both Chapters 2 and 4, a confirmation of the rare $b \to d$ and $b \to s$ processes requires a demonstration that the $\pi^+\pi^-$ final state originates from the $\rho^0(770)$ for the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ decay and from the $f_0(980)$ for the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ decay. In order to do this it is necessary to separate different states which contribute towards the $\pi^+\pi^-$ system using a dedicated analysis. This analysis would fit both the helicity angle distribution, cos(θ_h), and the invariant $m(\pi^+\pi^-)$ distributions with different resonance contributions which have different quantum numbers. Within this thesis a quick visual inspection of the cos(θ_h) and $m(\pi^+\pi^-)$ distributions is made. A visual inspection of the q^2 distributions is also made, where features such as the presence of a photon pole are indicative of a vector resonance (with $J > 0$) contribution due to the lack of a longitudinal photon polarisation within nature.

Figure 10.6 shows the $m(\pi^+\pi^-)$, q^2 and cos(θ_h) distributions from the data. These distributions are obtained by extracting a set of *sPlot* weights [166] from the final mass fits. These weights are then applied to the data distributions to unfold the signal before being compared to the simulated samples with the decay model weights applied. The simulated samples contain a mixture of states as was listed in Table 6.3 however are dominated by the $\rho^0(770)$ and $f_0(980)$ contributions.

The $m(\pi^+\pi^-)$ distributions from Figure 10.6 show that the agreement between the data and the $B_s^0 \to \pi^+\pi^-\mu^+\mu^-$ simulation is good. The main structure visible within the $\pi^+\pi^-$ spectrum is attributable to the $f_0(980)$ resonance. The equivalent plot for the B^0 mode shows the prominent structure from the $\rho^0(770)$ but the level of agreement between the data and simulation is worse. Specifically, it can be seen that the relative contributions of the different resonances to the $\pi^+\pi^-$ system differs between data and simulation. The data shows a larger contribution at higher invariant mass values above $1000\,\mathrm{MeV}/c^2$ compared to that predicted by simulation. This contribution may be attributable to the $f_2(1270)$ resonance. This suggests that the decay models from Chapter 6 do a good job of describing the B_s^0 decay but underestimate the contribution from the higher mass resonances for the B^0 mode. This conclusion can also be drawn from the equivalent q^2 plots where the level of agreement appears good for the B_s^0 mode showing that the $\pi^+\pi^-$ system is dominated by scalar resonance contributions. The B^0 mode shows poorer agreement between data and simulation in the q^2 region attributable to the photon pole. The data shows a larger contribution towards the photon pole than replicated in simulation suggesting that the $J > 0$ resonance contribution is underestimated. Future studies of the rare mode processes with more data could confirm the size of the contributing amplitudes within the $\pi^+\pi^-$ spectrum. This would allow for the construction of a more accurate decay model that would improve the agreement between data and simulation.

Only a limited discussion can be had regarding the helicity angle information due to the limited size of the data samples available to the analysis. The agreement seems broadly good between data and simulation however, without additional data the shape and corresponding quantum numbers cannot be clearly determined. In addition, the shapes seen are warped by the detector acceptance, by efficiencies and also by the interference between the various contributing resonances. This means that the naive shapes shown in Figure 4.2 cannot be easily identified. Within the B_s^0 mode distribution the helicity angle shows a drop off towards the +1 and -1 values of $\cos(\theta_h)$. This distribution is expected to be flat due to the major resonance contributions relevant to the B_s^0 mode being scalar in nature. There appears to be a slight discrepancy between data and simulation towards the -1 value. This is not believed to be a genuine effect as this angle should be symmetric and so this is expected to disappear in a future analysis with more data. The different shape between the B^0 and B_s^0 helicity angle distributions is caused by the larger vector resonance contribution into the B^0 mode.

10.3 Final Conclusions

To conclude, the analysis presented within this thesis succesfully provides a set of branching fraction measurements for the decays $B^0(B_s^0) \to \pi^+\pi^-\mu^+\mu^-$. The B_s^0 mode measurement in the full q^2 range is updated with the LHCb Run 2 data for the first time and the first confirmed observation of the B^0 mode in the full q^2 range is made using the combined Run 1 and Run 2 LHCb datasets. Furthermore, the analysis provides a statistically significant measurement of the branching fraction of both modes within the low q^2 region for the first time whilst providing an equivalent statistically significant measurement for the B^0 mode in the high q^2 region. To achieve a 5 σ confirmation of the B_s^0 mode within the high q^2 region, additional data will be required from future LHC runs. A larger dataset would also help to significantly reduce the size of the uncertainties on the experimental results as the largest systematic uncertainty relevant for the analysis is only around half the size of the statistical uncertainty. The results do not show any significant tensions with either the previous LHCb measurements nor with the theoretical predictions. The total uncertainty on the theoretical predictions is dominated by large form factor uncertainties and efforts within the theory community to reduce the size of these is of paramount importance to enable these results to be better compared to the SM predictions in the future.

The measurements performed make extensive use of a decay model to describe the different resonant contributions to the $\pi^+\pi^-$ final state. The accuracy of the models is checked after the completion of the analysis. The ability of the decay model to predict the distributions expected for the B_s^0 mode is seen to be very good. The B^0 mode decay model is seen to perform relatively well but systematically underestimates the vector resonance contributions to the system.

This is something which can be improved within a future analysis once the relative contributions to the $\pi^+\pi^-$ system are well known experimentally.

Confirmation of the rare mode requires the demonstration that the $\pi^+\pi^-$ pair originates from a $\rho^0(770)$ for the B^0 mode and a $f_0(980)$ for the B_s^0 mode. The data does indeed show large contributions from these two intermediate resonances. Nevertheless, a final confirmation of the rare modes requires the validation of the quantum numbers of the contributing resonances. This is something only possible with the inclusion of further data.

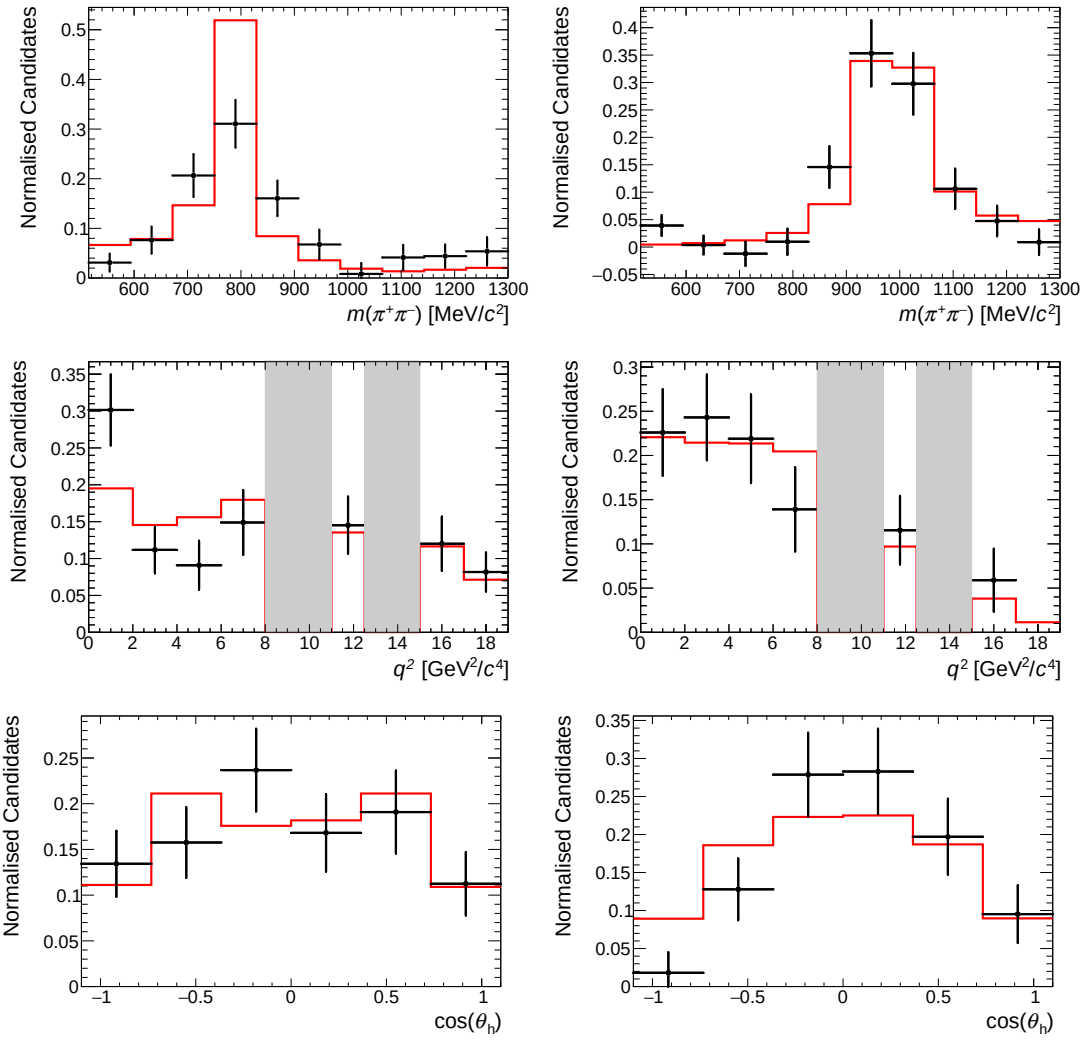

Figure 10.6: Comparisons between the unfolded signal data distributions (black) and equivalent simulated distributions (red) where the simulated samples have the decay model weights of Chapter 6 applied to include contributions from the resonances listed in Table 6.3. The left column applies to the B^0 mode while the right column to the B^0_s mode. The top row shows the $m(\pi^+\pi^-)$ distributions, the middle row the q^2 distributions, and the bottom row the distribution of the cosine of the pion helicity angle.

APPENDIX A

LHCb Simulation Software and Service Task

“ To err is human, but to really foul things up you need a computer ”

Paul R. Ehrlich,

This appendix provides a brief description of how Monté Carlo simulation is generated within the context of the LHCb data flow. In addition, a short discussion of service work completed for the collaboration is provided.

A.1 Monté Carlo Simulation and the LHCb data flow

Almost every analysis performed by LHCb, including the analysis presented within this thesis, makes extensive use of Monté-Carlo (MC) simulation samples. Simulated samples are used to evaluate efficiencies, model both signal shapes and yields and evaluate systematic uncertainties. LHCb MC samples are generated using a multi-step process with each step involving the application of software which fulfills a particular function. The entire generation process is encapsulated within LHCb's `Gauss` framework [196] and fits into the overall LHCb data flow, represented by the schematic shown in Figure A.1.

The first step in the `GAUSS` framework involves the generation of proton-proton collision events with the `PYTHIA` program [198] which initially generates a “hard” $gg \to q\bar{q}$ process. Then, the momenta of the different partons is determined using parton distribution functions

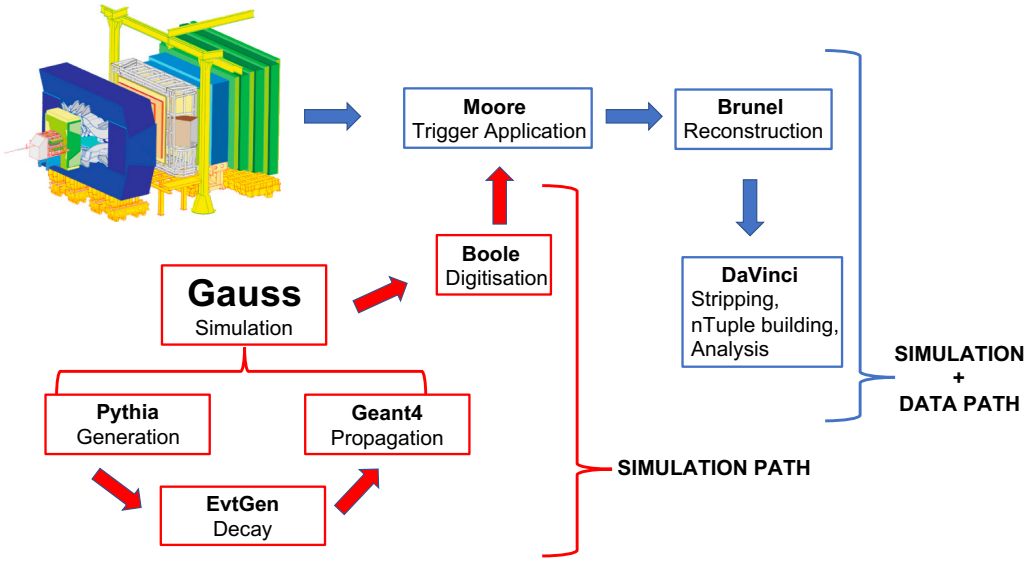

Figure A.1: The LHCb data flow with the path that the simulation takes in red before joining the same path as the data in blue. The LHCb drawing in the top left is taken from [197]

and this information is used to model the parton-parton interactions. The modelling of the parton interactions results in a realistic outgoing partonic structure which is then used to generate showers. These showers are subsequently organised into colour neutral "string" objects in order to fulfill the colour neutrality demanded by QCD. Finally, `PYTHIA` fragments the different string objects into hadrons. The LHCb-specific implementation of the software then repeats this process until a hadron emerges of the desired flavour which, for an LHCb analysis, is normally a beauty or charm hadron.

Once the hadronisation is completed the hadron needs to be decayed. The decay of heavy hadrons is achieved using the `C++ EvtGen` software package [199] which makes use of full amplitude models. Using amplitude information has the distinct advantage that the resulting models can describe complex physical phenomena, such as *CP* violation or angular correlations. `EvtGen` includes in excess of 70 different physics models and, due to it's modularity, it is easy for an analyst to write new physics models which can be merged into the code easily if the need arises. In the absence of a suitable decay model, `EvtGen` can be run with a simple phase space description which only takes into account kinematic information.

The final state particles which emerge after the application of `EvtGen` are then passed onto the next stage of the software chain, the `GEANT4` package [200]. This step includes the effect of the particle interactions with the materials of the LHCb detector itself and outputs the location of any charge deposits in sensitive sub-detectors. `GEANT4` includes electromagnetic,

hadronic and optical processes and is able to calculate the effect of these interactions in a wide variety of highly complex geometries.

The final output of GAUSS is passed through the Boole, Brunel and Moore software packages which simulate the digitisation of an electronic signal, the reconstruction of particle tracks and then both the software and hardware trigger decisions respectively. More information on these packages can be found in [201–203]. Finally, as is done with data, the simulated samples undergo a stripping selection and are converted into ROOT ntuples for off-line physics analysis using the DaVinci package [204]. In order to ensure that the simulated samples are as representative of reality as possible, the simulated samples are treated by the Brunel, Moore and DaVinci software packages almost identically to the way that the data is treated.

A.2 The DecFile package

For the LHCb analyst the interaction with EvtGen is primarily limited to the writing of a so-called DecFile [205]. A DecFile allows for the customisation of the EvtGen execution, providing a simulation sample tailored for a specific physics analysis. These customisations include forcing the mother particle to decay to a particular set of final state particles, via either a single or multiple decay trees. Branching fractions of the decay trees are either left as a parameter for the user to edit, or, by default, set to values taken from the Particle Data Group [154]. It is also possible within the DecFile for the user to overwrite physical parameters of the decay, such as the masses and widths of different particles. This can be useful for decays including particles where these properties are not currently well known or if the world values change with time. Users are also able to specify which decay model, if any, is required for the decay. The parameters of the model are specified as arguments to the model in the same order as they are in the EvtGen implementation.

Upon completion of a new DecFile the Gauss framework generates an options file which translates the decay in the DecFile into a format which is compatible with an execution of EvtGen. However, due to the large number of files used by the collaboration, and due to the extensive functionality of EvtGen, the supply of additional information which clarifies the nature of both the decay and the decay model is useful. This is achieved using a unique identifier for each file which comprises 8 separate flags – together known as the “Event Type”.

The different flags of the Event Type include:

- "**g**" The **general** flag – Indicates general information about the type of process, such as whether or not it contains a b quark
- "**s**" The **selection** flag – Indicates which particles are required to be present within the event, such as the assertion that a B_s^0 be included
- "**d**" The **decay** flag – Indicates whether the mother is forced to decay to a particular state, multiple states, or left to decay according to information from the Particle Data Group
- "**c**" The **charm and lepton** flag – Indicates how many charmonium, open charm, electrons or muons there are within the decay. This is necessary to determine whether a tertiary vertex from a charm hadron is expected and also to indicate the presence of particles relevant for the triggering process
- "**t**" The **track** flag – Indicates the number of charged tracks in the decay which are considered stable within the LHCb acceptance
- "**n**" The **neutral** flag – Indicates the number of neutral particles which can decay to final states, such as $\gamma\gamma$ which are relevant to the ECAL triggering decisions
- "**x**" The **extra** flag – Used to distinguish the `DecFile` from other, similar files. Also used to specify which resonance is to be produced if an excited beauty of charm resonance is requested
- "**u**" The **user** flag – Used to distinguish similar files when all other flags are the same as another `DecFile`

An example of an Event Type used within the analysis presented within this thesis is `11114022` which describes $B^0 \to \rho^0(770)\mu^+\mu^-$ decays where $\rho^0(770) \to \pi^+\pi^-$. Here, the general flag value of 1 indicates that the simulated sample contains a b quark. The selection flag is set to 1 to specifically select a B^0 mother while the decay flag is also set to 1 to force the decay of the $\rho^0(770)$ to a $\pi^+\pi^-$ final state. The charm and lepton flag is set to 1 to indicate the presence of muons within the final state and the track flag is set to 4 as there are four charged final state tracks. Finally, the neutral flag is set to 0 as their are no neutral particles in the forced part of the decay chain. The final two tracks are set to values to distinguish the decay from other, already existing configurations.

A second example is the Event Type `13144014` which is used within the analysis for the simulated sample of $B_s^0 \to J/\psi\, f_0(980)$ decays where $J/\psi \to \mu^+\mu^-$ and $f_0(980) \to \pi^+\pi^-$.

In this case the general flag is as before however the selection flag is set to a value of 3 to specify the B_s^0 mother. The decay flag is also set to the previous value whereas the charm and lepton flag takes the value of 4 to specify the existence of the charmonium J/ψ resonance. The remaining flags are set in the same way as previously.

Due to the reliance of LHCb physics analyses on simulated samples, the process of writing a sensible `DecFile` is a crucial step within the analysis procedure. As part of my service task for the LHCb collaboration I helped to maintain the `DecFile` software package. This involved checking all new `DecFiles` written by analysts and running tests on them to ensure that they successfully generated test samples. These tests included making sure that the CPU time for the generation was appropriate, ensuring that the content of the `DecFile` was compatible with the chosen Event Type, ensuring that particular selection and acceptance cuts were as intended for the analysis, and ensuring that the documentation provided by the proponent was complete and comprehensive. Once a `DecFile` was complete and successfully passed all tests, I would merge the new file with the existing software package. Over the course of my service task several hundred new files were added to the software package.

APPENDIX B

Invariant Mass Fit Parameters

“ If you have a procedure with 10 parameters, you probably missed some. ”

Alan Perlis,

This appendix includes parameter values, not included within the main body of this thesis, resulting from the invariant mass fits to the MC and data samples used within the analysis. Fit parameters relevant for the control mode include the double Crystal Ball + Gaussian fit to the MC $B^0 \to J/\psi K^{*0}$ mode which can be found in Table B.1 and the two separate Gaussian and single Crystal Ball fits to data to extract the fit parameters for the $B^0 \to J/\psi K^+$ mode which can be found in Table B.2. Fit parameters relevant for the resonant mode include the double Crystal Ball + Gaussian fit to the MC $B^0 \to J/\psi \pi^+\pi^-$ mode found in Table B.3; the double Crystal Ball fit to the MC $B^0 \to J/\psi K^+\pi^-$ mode found in Table B.4; the double Crystal Ball fit to the MC $B^0_s \to J/\psi \eta'$ mode found in Table B.5 and the Bifurcated Gaussian + Gaussian fit to the `RapidSim` B^+_c mode found in Table B.6. Fit parameters relevant for the rare mode include the double Crystal Ball fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ mode within the data found in Table B.7; the double Crystal Ball + Gaussian fit to the MC $B^0 \to \pi^+\pi^-\mu^+\mu^-$ mode found in Table B.3 and the double Crystal Ball fit to the MC $B^0 \to K^+\pi^-\mu^+\mu^-$ mode found in Table B.9.

Table B.1: Parameters from the fit to the $B^0 \to J/\psi K^{*0}$ MC to determine the signal lineshape for the control mode invariant mass fits. The parameters are shown for 2011+2012, 2015+2016 and 2017+2018.

Parameter	2011+2012	2015+2016	2017+2018
α_1	-2.13 ± 0.06	-2.10 ± 0.23	-2.00 ± 0.20
α_2	0.41 ± 0.02	0.44 ± 0.07	0.47 ± 0.06
n_1	3.27 ± 0.14	3.26 ± 0.32	3.41 ± 0.29
f_1	0.74 ± 0.04	0.69 ± 0.19	0.62 ± 0.16
f_{G}	0.64 ± 0.10	0.54 ± 0.28	0.48 ± 0.18
σ_{CB}	17.35 ± 0.68	18.37 ± 2.37	18.75 ± 1.74
σ_{G}	0.67 ± 0.01	0.69 ± 0.01	0.69 ± 0.02
μ_{B^0}	5280.70 ± 0.06	5281.12 ± 0.16	5281.13 ± 0.12

Fit	Parameter	2011+2012	2015+2016	2017+2018
$m(J/\psi K^+\pi^-)$ Shape fit	α_{CB}	-0.44 ± 0.08	-0.50 ± 0.10	-0.47 ± 0.08
	μ_{B^+}	5480.65 ± 9.19	5481.72 ± 10.24	5471.90 ± 8.70
	σ_{CB}	41.90 ± 6.12	46.43 ± 7.04	47.11 ± 5.88
	λ	-0.01 ± 0.02	-0.01 ± 0.02	-0.01 ± 0.02
	$N(B^+ \to J/\psi K^+)$	246.00 ± 15.37	165.00 ± 12.47	301.02 ± 17.07
	N_{bkg}	0.000 ± 0.002	0.000 ± 0.000	0.000 ± 0.001
$m(J/\psi K^+)$ Yield fit	μ_{B^+}	5282.77 ± 1.323	5280.43 ± 1.61	5281.24 ± 1.40
	σ_{G}	17.49 ± 1.10	16.61 ± 1.29	18.23 ± 1.14
	λ	-0.012 ± 0.002	-0.014 ± 0.002	-0.010 ± 0.001
	$N(B^+ \to J/\psi K^+)$	220 ± 16	136 ± 13	232 ± 17
	N_{bkg}	48 ± 9	56 ± 9	114 ± 13

Table B.2: Parameters from the fits to the invariant $J/\psi K^+\pi^-$ and $J/\psi K^+$ mass distributions to determine the shape and yield of the $B^+ \to J/\psi K^+$ background within the control mode invariant mass fits.

Table B.3: Parameters from the fit to the $B^0 \to J/\psi\,\pi^+\pi^-$ MC for 2011+2012, 2015+2016 and 2017+2018 with and without the J/ψ mass constraint. The parameters are used within the final resonant mode invariant mass fits.

Fit	Parameter	2011+2012	2015+2016	2017+2018
No J/ψ Constraint	α_1	-2.39 ± 0.17	-2.37 ± 0.17	-2.14 ± 0.51
	α_2	0.41 ± 0.10	0.50 ± 0.15	0.40 ± 0.10
	f_1	0.74 ± 0.17	0.69 ± 0.15	0.70 ± 0.32
	f_2	0.69 ± 0.40	0.71 ± 0.21	0.66 ± 0.64
	σ_{CB}	17.78 ± 2.49	18.00 ± 1.74	17.23 ± 3.09
	$\sigma_{\mathrm{G}}/\sigma_{\mathrm{CB}}$	0.66 ± 0.11	0.59 ± 0.06	0.73 ± 0.14
	μ_{B^0}	5280.76 ± 0.22	5281.18 ± 0.23	5281.39 ± 0.30
J/ψ Constraint	α_1	-1.21 ± 0.34	-1.55 ± 0.60	-0.70 ± 0.79
	α_2	0.40 ± 0.11	0.43 ± 0.16	0.55 ± 0.14
	f_1	0.68 ± 0.15	0.75 ± 0.29	0.41 ± 0.34
	f_2	0.25 ± 0.06	0.45 ± 0.51	0.24 ± 0.09
	σ_{CB}	12.60 ± 1.56	10.39 ± 4.52	11.60 ± 3.28
	$\sigma_{\mathrm{G}}/\sigma_{\mathrm{CB}}$	0.53 ± 0.06	0.59 ± 0.10	0.60 ± 0.16
	μ_{B^0}	5279.97 ± 0.05	5280.48 ± 0.13	5280.33 ± 0.10

Fit	Parameter	2011+2012	2015+2016	2017+2018
No J/ψ Constraint	α_1	-1.84 ± 0.04	-0.14 ± 1.18	-1.58 ± 4.72
	α_2	0.001 ± 0.772	0.715 ± 0.421	0.000 ± 0.010
	n_1	0.02 ± 65.24	0.02 ± 0.20	0.01 ± 0.01
	f_1	0.33 ± 0.65	0.00 ± 0.01	0.00 ± 0.00
	σ_{CB}	22.75 ± 1.44	23.92 ± 7.16	35.62 ± 10.39
	μ_{B^0}	5219.34 ± 12.32	5214.42 ± 10.07	5198.62 ± 17.06
J/ψ Constraint	α_1	-0.52 ± 0.13	-0.57 ± 0.85	-1.13 ± 1.92
	α_2	0.14 ± 0.06	0.33 ± 3.09	0.14 ± 1.69
	n_1	9.99 ± 18.31	0.74 ± 13.60	0.01 ± 0.01
	f_1	0.06 ± 0.02	0.05 ± 0.11	0.13 ± 0.56
	σ_{CB}	12.09 ± 0.89	11.50 ± 9.65	19.51 ± 10.99
	μ_{B^0}	5226.74 ± 1.53	5222.61 ± 17.49	5218.02 ± 13.87

Table B.4: Parameters from the fit to the $B^0 \to J/\psi\,K^+\pi^-$ MC for 2011+2012, 2015+2016 and 2017+2018 for the case with and without a J/ψ mass constraint. The parameters are used within the final resonant mode invariant mass fits.

Fit	Parameter	2011+2012	2015+2016	2017+2018
No J/ψ Constraint	α_1	-0.40 ± 0.24	-0.72 ± 0.06	-0.53 ± 0.06
	α_2	0.11 ± 0.02	0.08 ± 0.01	0.08 ± 0.01
	f_1	0.001 ± 0.001	0.007 ± 0.001	0.004 ± 0.001
	σ_{CB}	35.00 ± 1.30	33.07 ± 0.57	33.82 ± 0.42
	μ_{B^0}	5281.11 ± 2.49	5284.93 ± 1.05	5283.84 ± 0.79
J/ψ Constraint	α_1	-0.73 ± 0.10	-0.77 ± 0.03	-0.62 ± 0.05
	α_2	0.10 ± 0.02	0.07 ± 0.01	0.08 ± 0.00
	f_1	0.010 ± 0.000	0.020 ± 0.000	0.010 ± 0.000
	σ_{CB}	30.26 ± 1.62	28.16 ± 0.57	28.86 ± 0.41
	μ_{B^0}	5286.18 ± 2.90	5289.12 ± 0.96	5288.52 ± 0.71

Table B.5: Parameters from the fit to the $B_s^0 \to J/\psi\,\eta'$ MC for 2011+2012, 2015+2016 and 2017+2018 with and without the J/ψ mass constraint. The parameters are used within the final resonant mode mass fits.

Table B.6: Parameters from the fit to the B_c^+ `RapidSim` sample. The L and R subscripts denote the left and right components of the bifurcated Gaussian which correspond to side of the Gaussian lower than, and higher than the mean respectively. The parameters are used within the final resonant mode invariant mass fits.

Parameter	Value
f_{G}	0.38 ± 0.01
μ_{G}	4800.00 ± 4.02
μ_{BG}	6076.69 ± 1.66
σ_{G}	347.11 ± 7.10
σ_{L}	500.53 ± 17.71
σ_{R}	32.17 ± 0.96

Dataset	Parameter	Full q^2	Low q^2	High q^2
2011+2012	μ_{CB}	5284.94 ± 0.76	5284.64 ± 0.92	5285.22 ± 1.97
	σ_{CB}	16.7 ± 0.8	15.7 ± 1.0	17.8 ± 2.0
	a_{CB1}	-1.65 ± 0.34	-1.53 ± 0.36	-0.16 ± 0.29
	a_{CB2}	0.85 ± 0.26	0.65 ± 0.45	1.00 ± 0.23
	n_{CB1}	0.78 ± 0.55	1.58 ± 0.91	2.56 ± 0.92
	n_{CB2}	0.85 ± 0.26	0.65 ± 0.45	1.00 ± 0.23
	f_{CB1}	0.36 ± 0.24	0.52 ± 0.26	0.13 ± 0.09
	λ	-0.029 ± 0.041	-0.050 ± 0.013	-0.050 ± 0.001
	N_{bkg}	30 ± 38	8 ± 13	2 ± 2
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	1202 ± 51	840 ± 31	218 ± 15
2015+2016	μ_{CB}	5278.80 ± 0.69	5278.49 ± 0.84	5279.88 ± 0.57
	σ_{CB}	16.6 ± 0.7	16.3 ± 1.0	15.7 ± 0.6
	a_{CB1}	-0.39 ± 0.28	-1.36 ± 0.56	-1.47 ± 0.29
	a_{CB2}	1.41 ± 0.26	1.28 ± 0.55	0.97 ± 0.28
	n_{CB1}	1.47 ± 0.87	1.46 ± 1.00	2.40 ± 1.32
	n_{CB2}	1.83 ± 0.76	1.89 ± 1.42	1.42 ± 1.13
	f_{CB1}	0.13 ± 0.12	0.40 ± 0.36	0.47 ± 0.26
	λ	-0.018 ± 0.042	-0.050 ± 0.003	-0.012 ± 0.009
	N_{bkg}	49 ± 67	7 ± 9	0 ± 1
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	1188 ± 115	856 ± 30	202 ± 24
2017+2018	μ_{CB}	5280.33 ± 0.51	5279.88 ± 0.57	5282.76 ± 1.43
	σ_{CB}	15.9 ± 0.5	15.7 ± 0.6	15.3 ± 1.2
	a_{CB1}	-1.48 ± 0.31	-1.47 ± 0.29	-5.92 ± 4.91
	a_{CB2}	1.02 ± 0.21	0.97 ± 0.28	0.97 ± 0.18
	n_{CB1}	1.32 ± 0.77	2.40 ± 1.38	2.28 ± 1.22
	n_{CB2}	2.45 ± 1.27	2.61 ± 0.45	4.21 ± 2.82
	f_{CB1}	0.31 ± 0.19	0.47 ± 0.26	0.02 ± 0.13
	λ	-0.014 ± 0.012	-0.012 ± 0.009	-0.004 ± 0.004
	N_{bkg}	166 ± 103	154 ± 83	25 ± 9
	$N(B^0 \to K^+\pi^-\mu^+\mu^-)$	2479 ± 114	1707 ± 92	392 ± 21

Table B.7: Parameters from the fit to the $m(K^+\pi^-\mu^+\mu^-)$ distribution for the different datasets and the different q^2 regimes. The yield is scaled using efficiencies and used to constrain the $B^0 \to K^+\pi^-\mu^+\mu^-$ background within the final rare mode invariant mass fits.

Table B.8: Parameters from the fit to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC for 2011+2012, 2015+2016 and 2017+2018. The parameters are used within the final rare mode invariant mass fits.

Parameter	2011+2012	2015+2016	2017+2018
α_1	-1.57 ± 0.13	-0.62 ± 0.05	-2.15 ± 0.04
α_2	1.16 ± 0.06	1.61 ± 0.07	0.64 ± 0.09
n_1	3.76 ± 0.21	6.80 ± 0.79	3.14 ± 0.19
n_2	1.56 ± 0.08	1.35 ± 0.14	1.71 ± 0.56
f_1	0.44 ± 0.05	0.06 ± 0.02	0.79 ± 0.00
f_2	0.69 ± 0.10	0.73 ± 0.16	0.82 ± 0.05
μ_{B^0}	5280.48 ± 0.06	5280.56 ± 0.09	5281.10 ± 0.09
σ_{CB}	13.94 ± 0.18	14.05 ± 0.13	14.76 ± 0.20
$\sigma_{\mathrm{G}}/\sigma_{dCB}$	1.33 ± 0.08	1.32 ± 0.26	1.43 ± 0.02

Table B.9: Parameters from the fit to the $B^0 \to K\pi\mu^+\mu^-$ MC fit for 2011+2012, 2015+2016 and 2017+2018. The parameters are used within the final rare mode invariant mass fits.

Parameter	2011+2012	2015+2016	2017+2018
α_1	-0.63 ± 0.67	-2.37 ± 3.58	-0.03 ± 0.53
α_2	0.25 ± 0.20	7.39 ± 2.40	0.27 ± 0.02
f_1	0.021 ± 0.099	0.270 ± 0.835	0.001 ± 0.024
σ_1	25.14 ± 6.37	32.36 ± 6.31	15.78 ± 9.72
σ_2	20.70 ± 1.82	32.64 ± 2.69	22.17 ± 0.93
μ_{B^0}	5227.59 ± 3.17	5204.17 ± 2.87	5224.00 ± 0.44

APPENDIX C

Additional Figures

“ Mankind invented a system to cope with the fact that we are so intrinsically lousy at manipulating numbers. It’s called the graph. ”

Charlie Munger,

This appendix contains additional material not included within the main body of the thesis. These include figures which add additional details to the trigger efficiency correction discussion from Chapter 7 and versions of the invariant mass fits from Chapter 8 where the goodness of fit is included via residual pull plots.

C.1 Trigger Efficiency Correction - Additional Material

The L0 trigger efficiency corrections are determined in Chapter 7 using the TISTOS method. The final correcting weights are shown in Figure 7.3 where the values of the correcting weights are shown to be approximately unity. From Figure 7.3 it can be seen that some adjacent bins within the 2D histograms show quite different values to one another. This is particularly seen in the largest bin of maximum muon p_{T}. This effect results from the binning scheme used whereby certain bins cover areas of phase space with a relatively small number of candidates meaning that a fine binning scheme can result in noticeable efficiency jumps between adjacent bins. This effect can be seen in the equivalent plots containing the MC L0 trigger efficiency (Figure C.1) and data L0 trigger efficiency (Figure C.2) values which together are used to determine the correction weights. This effect could be removed by absorbing some of the smaller bins in this region of

phase space into larger bins which would subsequently reduce the difference between adjacent bin values. However, as the L0 trigger correction values are almost unity, and because the corrections largely cancel within the ratio between the rare mode and control mode measurements, this effect is small enough to be ignored.

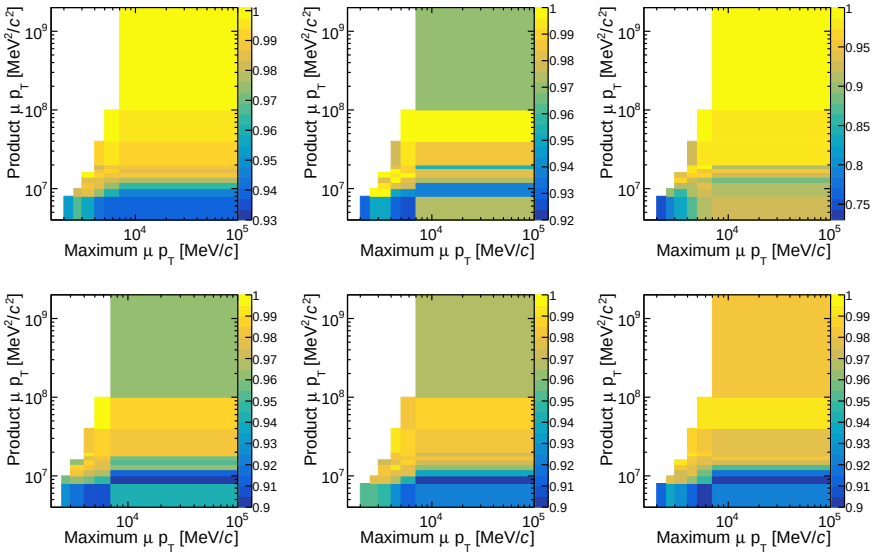

Figure C.1: The L0 trigger efficiency evaluated from the MC samples in bins of maximum muon p_{T} and the product of muon p_{T} values. The figures show, from the left to right on the top row the 2011, 2012 and 2015 histograms. On the bottom row the equivalent histograms are shown for the 2016, 2017 and 2018 samples.

C.2 Invariant Mass Fits - Goodness of Fit

This section presents the invariant mass fits from Chapter 8 with additional pull plots. The pull plots show the difference between the fitted shape and the data/MC, divided by the corresponding uncertainty. The pulls therefore provide a visual way of determining the goodness of fit. The pull plots are not included within the main body of the thesis to conserve space, however are presented here for completeness.

Figure C.3 shows the fit to the $B^0 \to J/\psi\, K^{*0}$ MC samples to model the control mode signal shape, Figure C.4 shows the fits to the $B^+ \to J/\psi\, K^+$ mode within the data and Figure C.5 shows the final control mode fits.

For the resonant mode Figure C.6 shows the fits to the $B^0 \to J/\psi\, \pi^+\pi^-$ MC, Figure C.7 shows the fits to the $B^0 \to J/\psi\, K^+\pi^-$ MC, Figure C.8 shows the fits to the $B^0_s \to J/\psi\, \eta'$ MC,

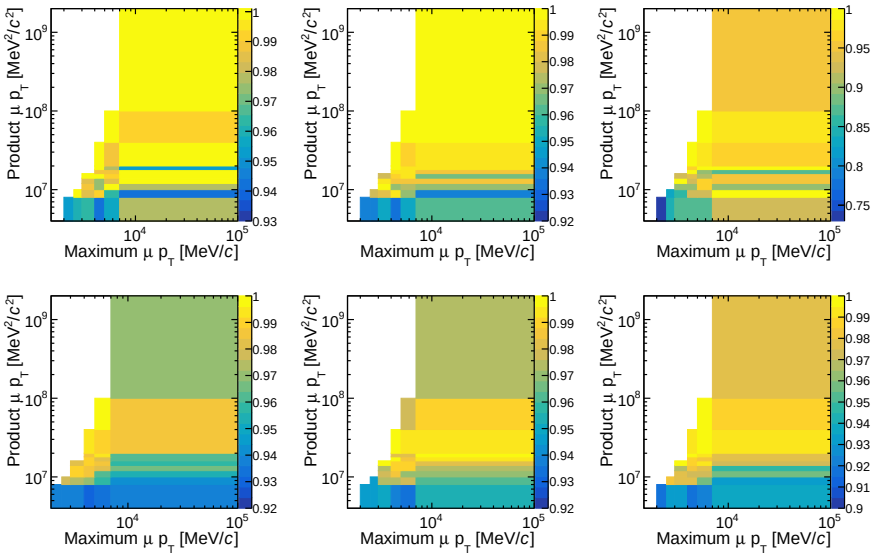

Figure C.2: The L0 trigger efficiency evaluated from the data samples in bins of maximum muon p_{T} and the product of muon p_{T} values using the TISTOS method. The figures show, from the left to right on the top row the 2011, 2012 and 2015 histograms. On the bottom row the equivalent histograms are shown for the 2016, 2017 and 2018 samples.

Figure C.9 shows the fits to the $B^0_s \to J/\psi\,\phi$ MC, Figure C.10 shows the fit to the B^+_c background `RapidSim` sample, Figure C.11 shows the final resonant mode fits and Figure C.12 shows the final resonant mode fits without the background yield constraints applied.

Finally, Figure C.13 shows the fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC, Figure C.14 shows the fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC, Figure C.15 shows the final rare mode fits for the full q^2 range, Figure C.16 shows the final rare mode fits for the low q^2 range, Figure C.17 shows the final rare mode fits for the high q^2 range and finally, Figure C.18 shows the final rare mode fits performed separately on the Run 1 and two Run 2 datasets.

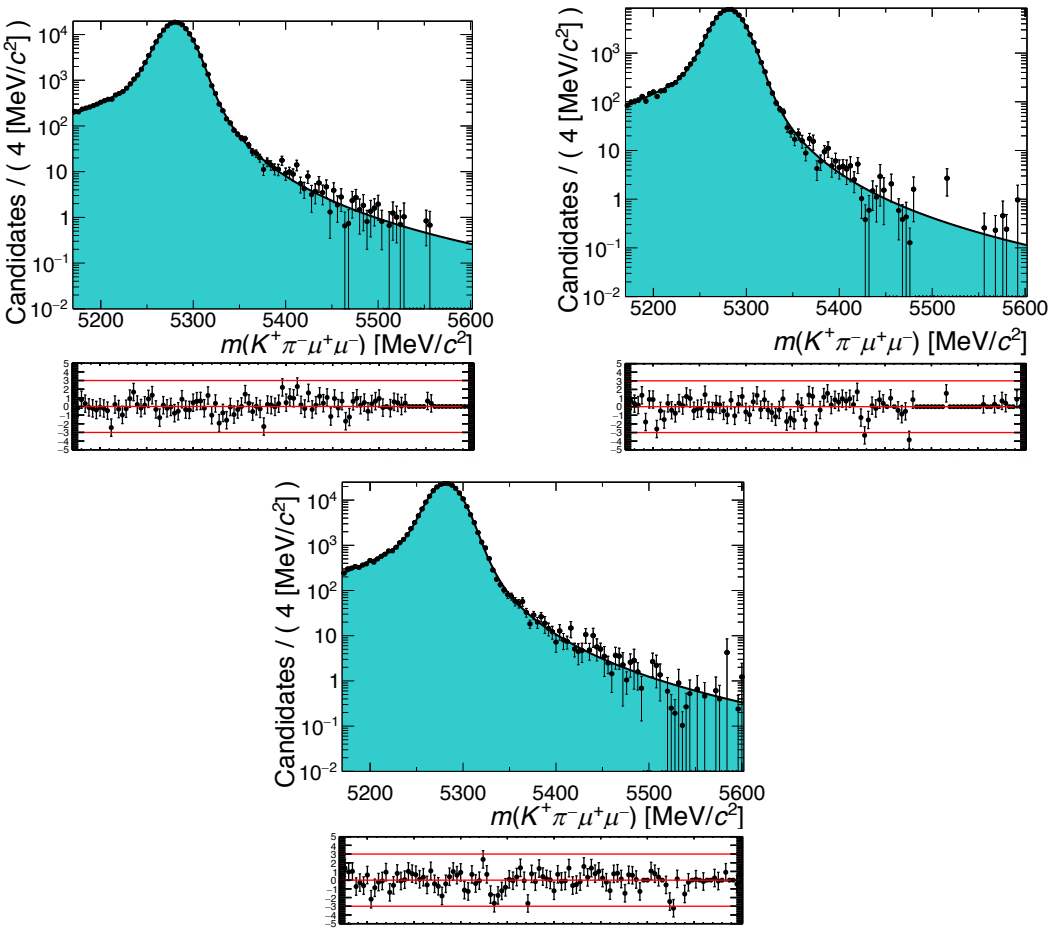

Figure C.3: Invariant mass fits to the $B^0 \to J/\psi\, K^{*0}$ MC samples using the sum of two Crystal Ball functions and a Gaussian function for the 11+12 (top left), 15+16 (top right) and 17+18 (bottom) MC samples. The candidates are weighted with the full set of analysis correcting weights.

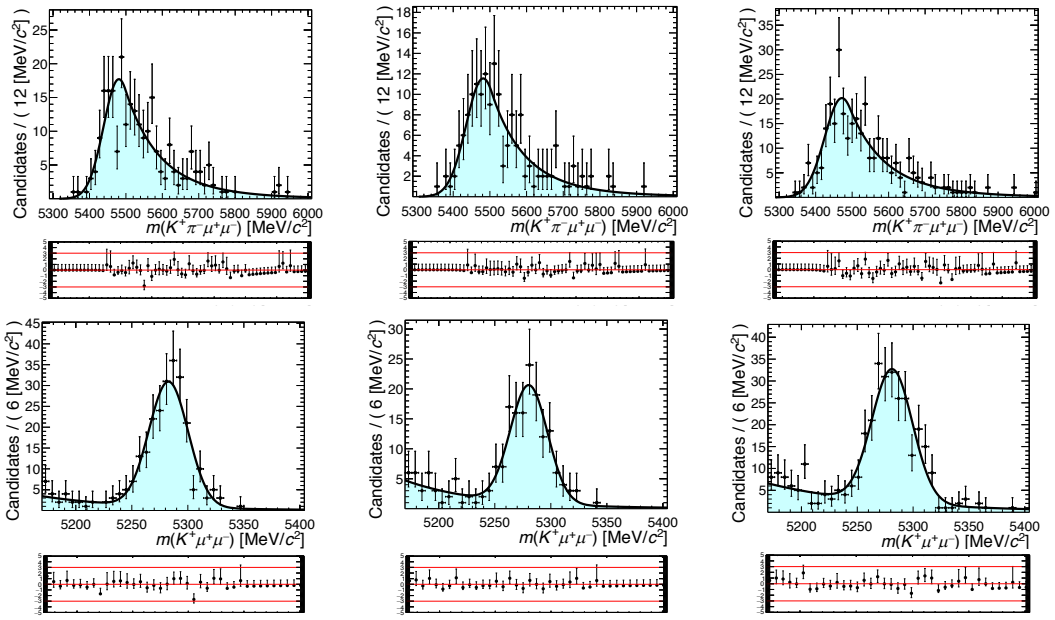

Figure C.4: Invariant mass fits to the three datasets to determine the shape and yield of $B^+ \to J/\psi\, K^+$ decays present within the final control mode fit. The top row shows the fit to the $m(K^+\pi^-\mu^+\mu^-)$ invariant mass using a single CB shape for the 2011+2012 (left), 2015+2016 (middle) and 2017+2018 (right) datasets to extract a background shape. The bottom row shows, for the same three datasets, Gaussian fits to the $m(K^+\mu^+\mu^-)$ distribution to extract an estimated background yield.

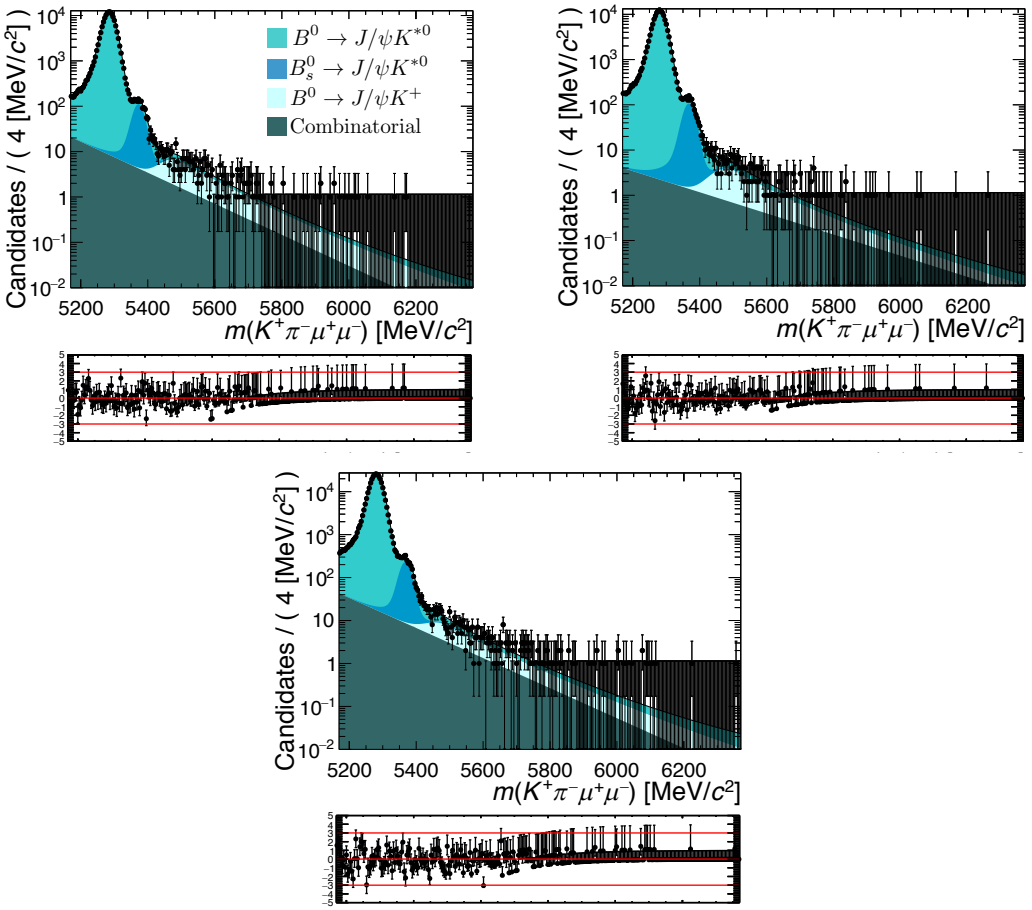

Figure C.5: The invariant mass fits to the $B^0 \to J/\psi K^{*0}$ control mode for the 2011+2012 dataset (top left), the 2015+2016 dataset (top right) and the 2017+2018 dataset (bottom).

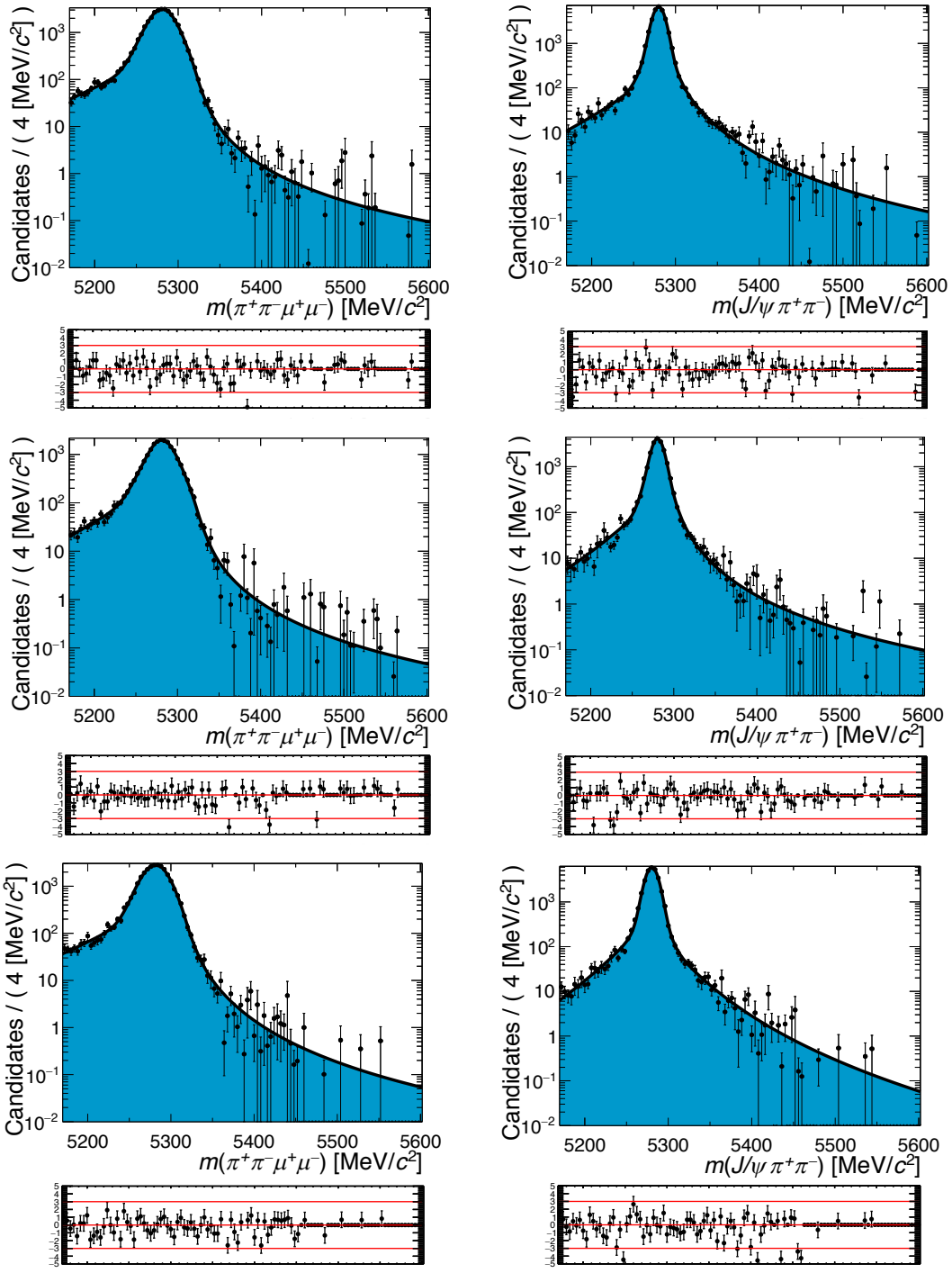

Figure C.6: Invariant mass fits to the $B^0 \to J/\psi\,\pi^+\pi^-$ MC samples which are used to determine the signal shapes for the final resonant mode fits. The figures are the same as shown in Figure 8.5 with pull plots included.

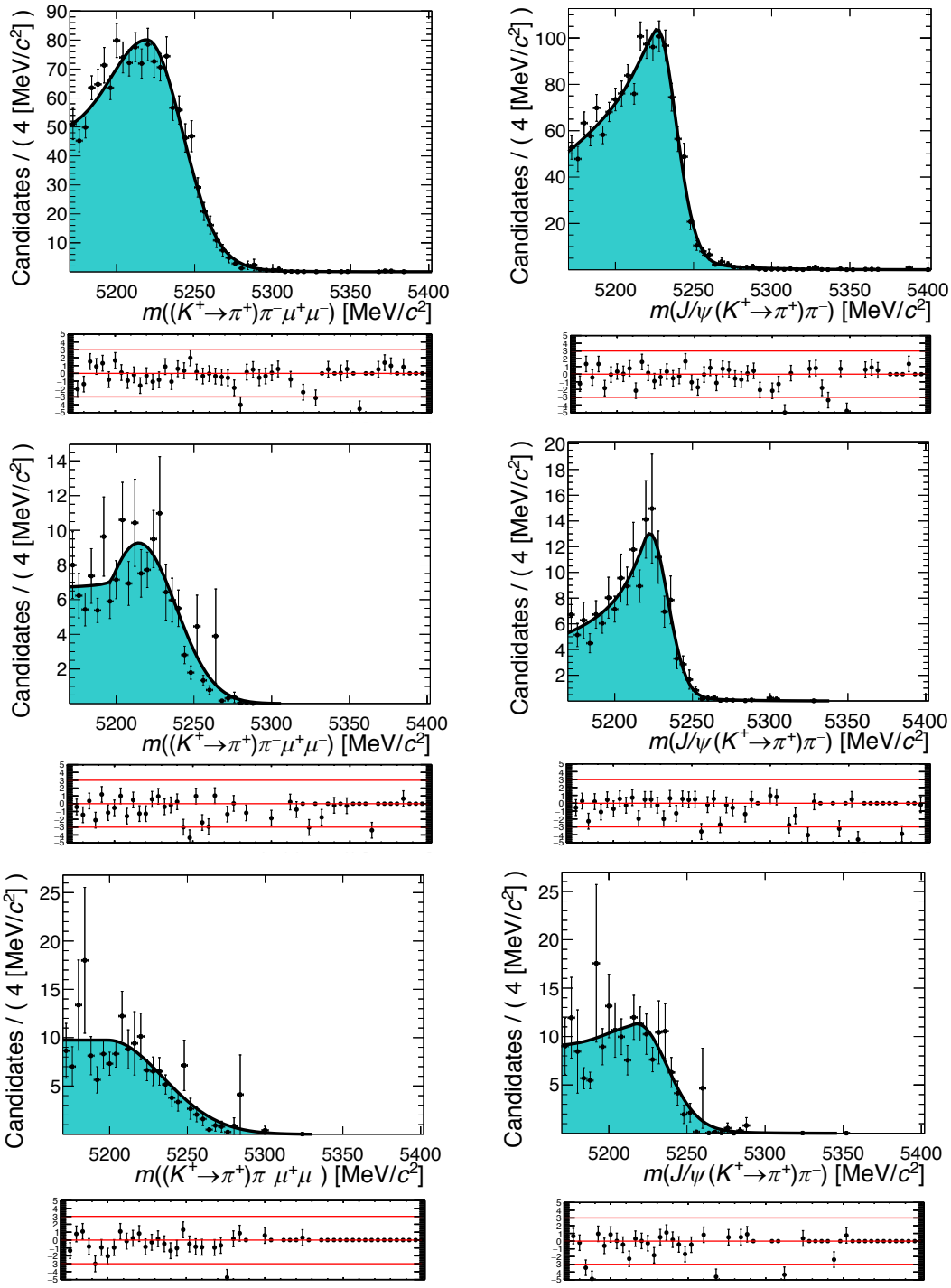

Figure C.7: Invariant mass fits to the $B^0 \to J/\psi\, K^+\pi^-$ MC samples made with the sum of two CB shapes. The figures are the same as shown in Figure 8.6 with added pull plots.

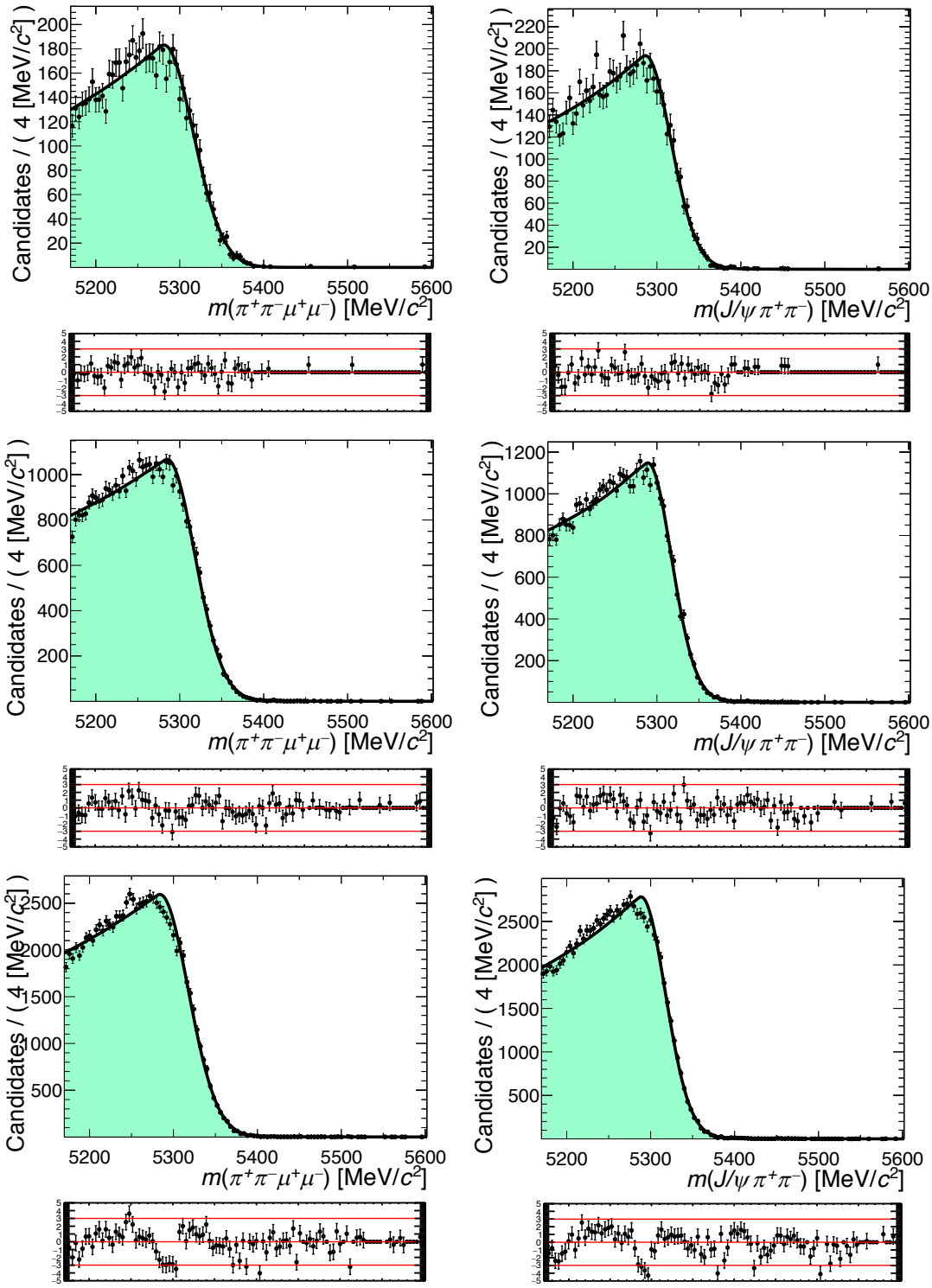

Figure C.8: Invariant mass fits to the $B_s^0 \to J/\psi\,\eta'$ MC samples made with the sum of two CB shapes. The plots are the same as shown in Figure 8.7 with added pull plots.

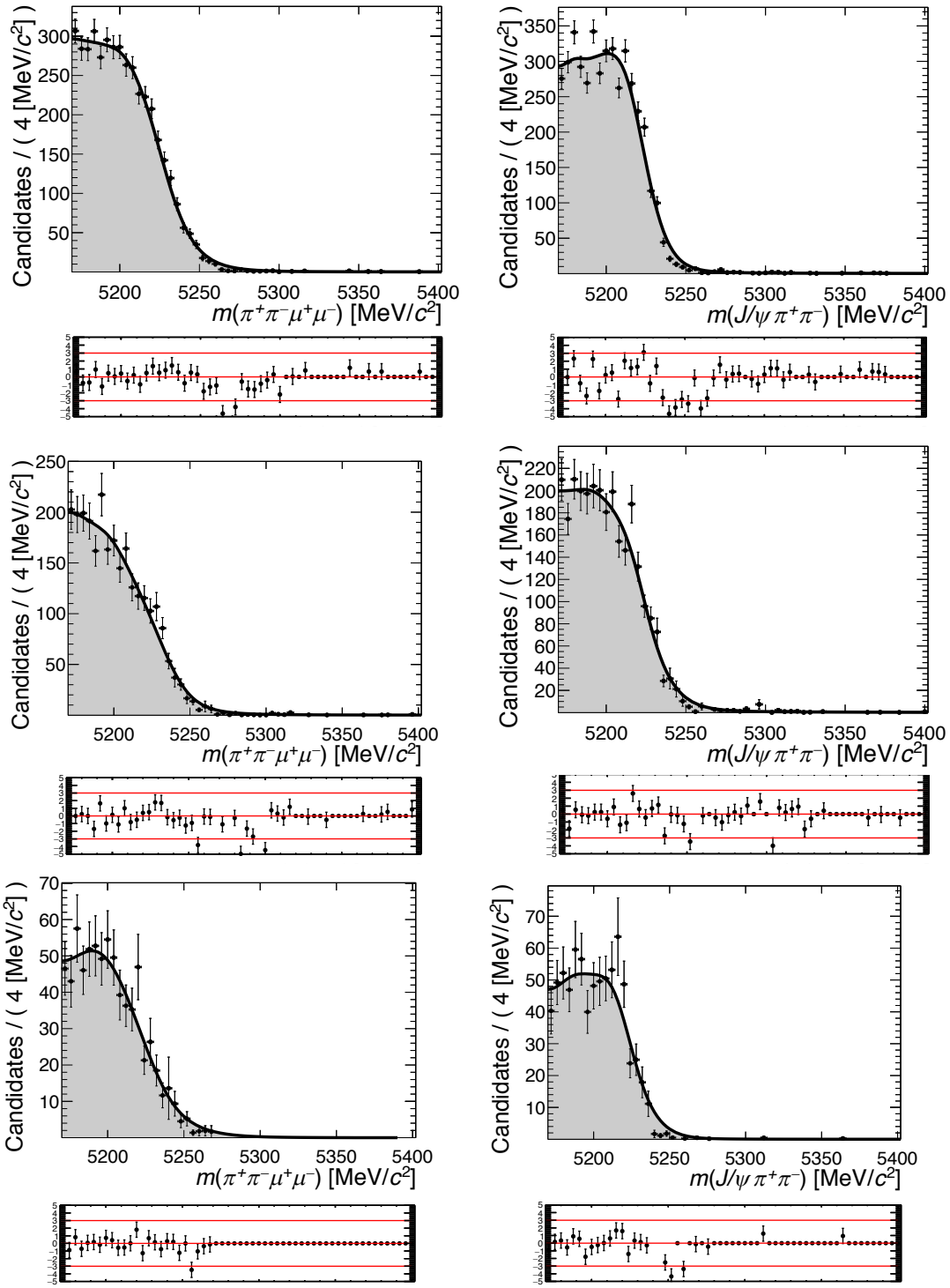

Figure C.9: Invariant mass fits to the $B_s^0 \to J/\psi\,\phi$ MC samples implemented via a KDE. The fits are the same as shown in Figure 8.8 with added pull plots.

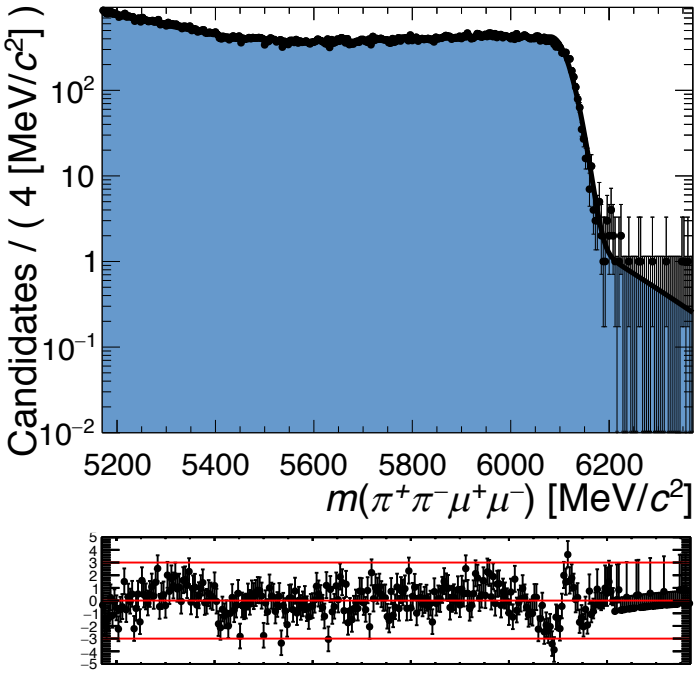

Figure C.10: The invariant mass fit to the `RapidSim` simulated sample of B_c^+ background.

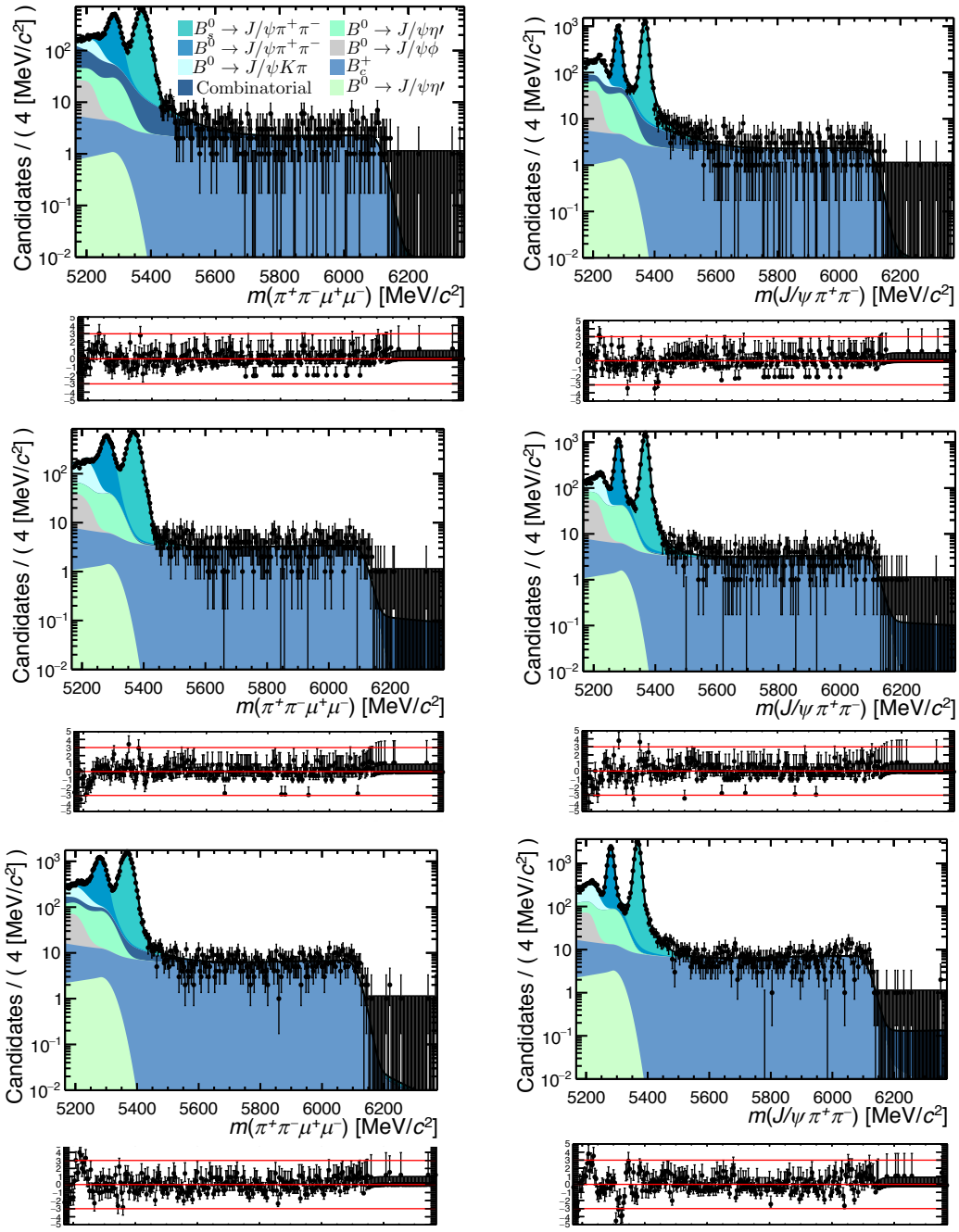

Figure C.11: The final resonant mode invariant mass fits for the 2011+2012 (top), 2015+2016 (middle) and 2017+2018 (bottom) datasets on a logarithmic scale to show the different background components. The left column shows the fits without the J/ψ mass constraint applied and the right plots with the J/ψ mass constraint applied.

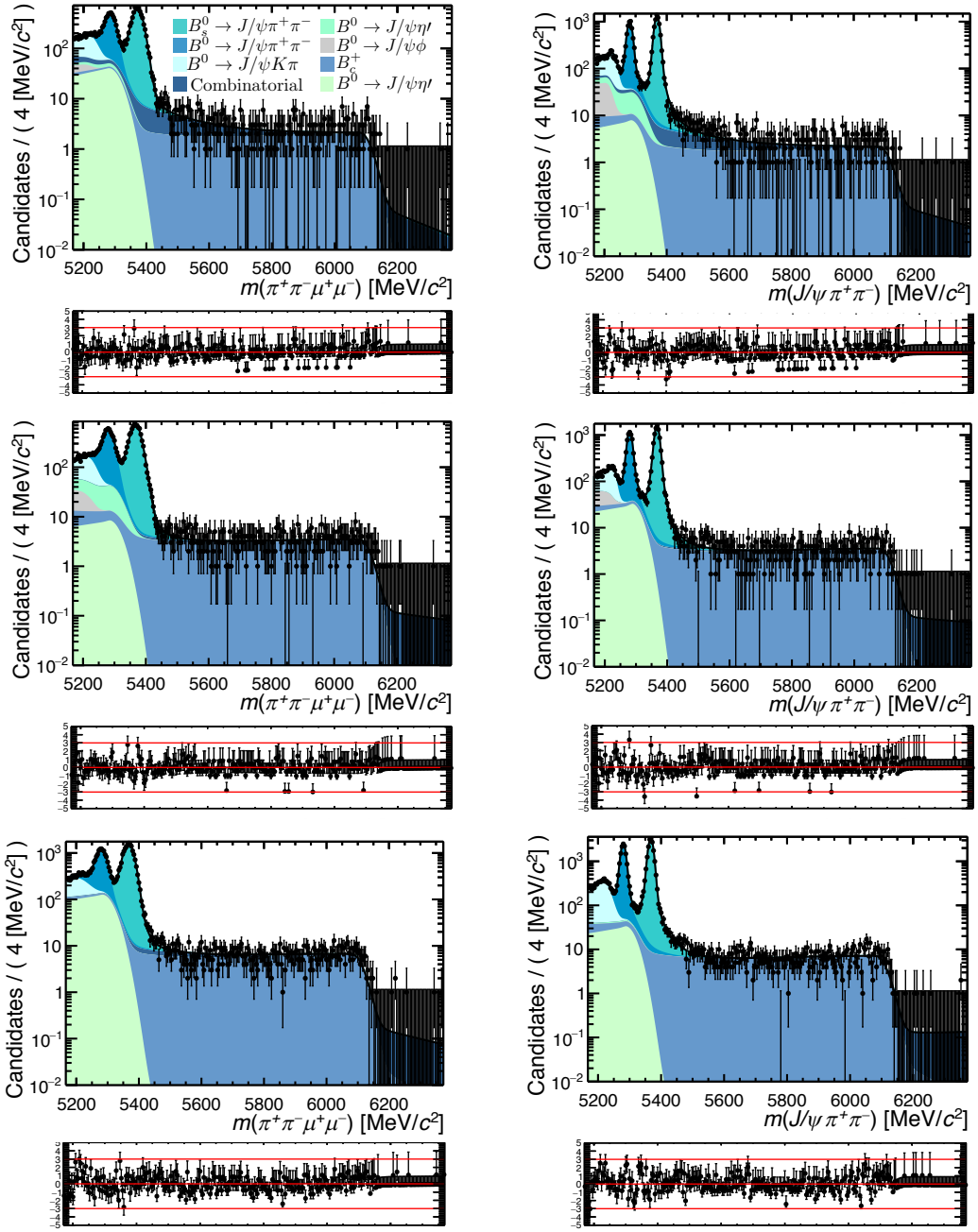

Figure C.12: The final resonant mode invariant mass fits for the 2011+2012 (top), 2015+2016 (middle) and 2017+2018 (bottom) datasets on a logarithmic scale to show the different background components where no yield constraints are applied. The left column shows the fits without the J/ψ mass constraint applied and the right plots with the J/ψ mass constraint applied.

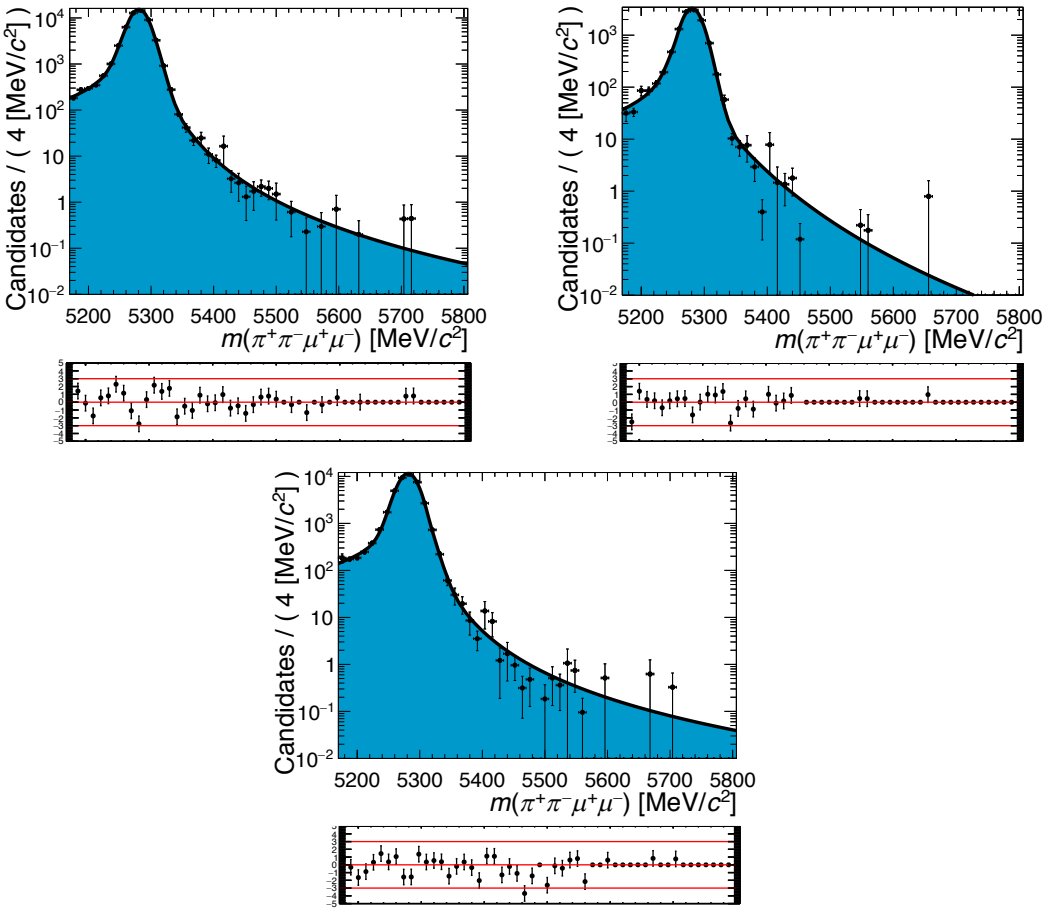

Figure C.13: Invariant mass fits to the $B^0 \to \pi^+\pi^-\mu^+\mu^-$ MC samples which serves as a proxy for the signal shapes within the rare mode fits. The fits are performed with the sum of two CB shapes and a Gaussian shape and can be seen for the 2011+2012 sample (top left), the 2015+2016 sample (top right) and the 2017+2018 sample (bottom). The candidates are weighted with the full set of analysis correcting weights.

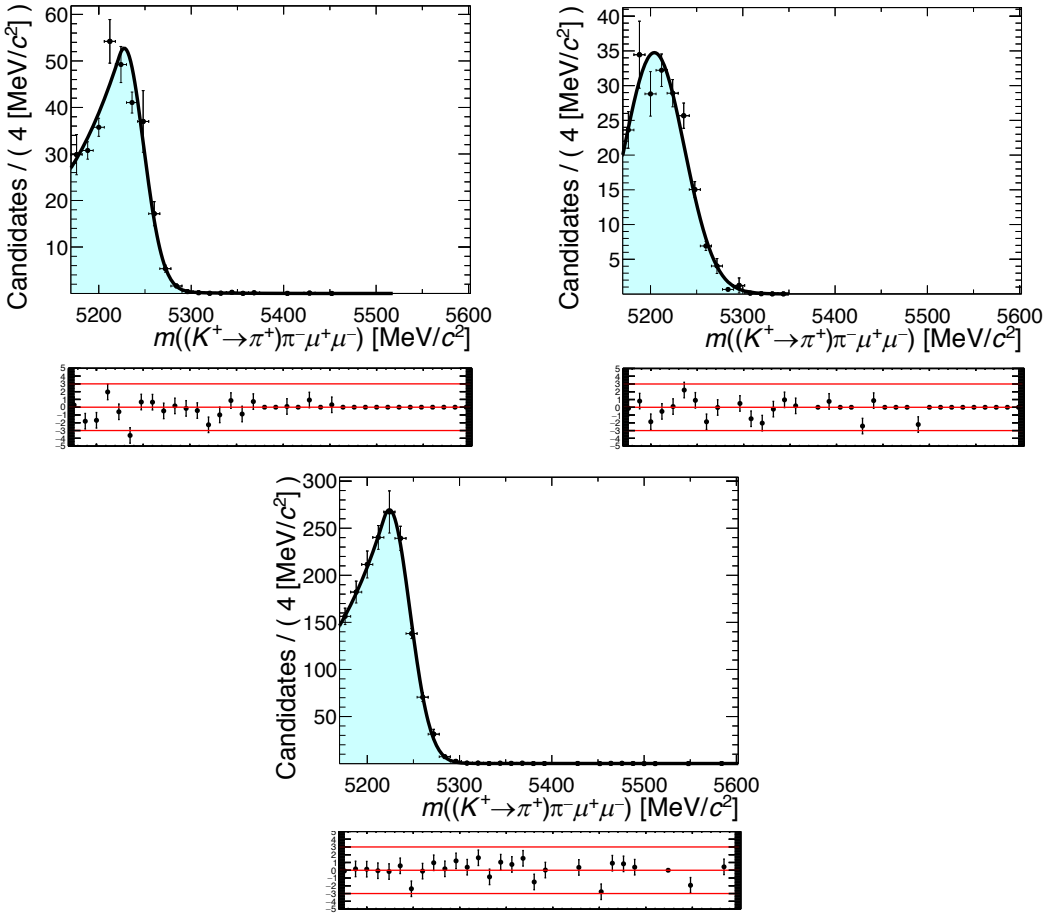

Figure C.14: The invariant mass fits to the $B^0 \to K^+\pi^-\mu^+\mu^-$ MC samples for the 2011+2012 (top left), 2015+2016 (top right) and the 2017+2018 sample (bottom). The fits are performed using the sum of two CB shapes. The candidates are weighted with the full set of analysis correcting weights.

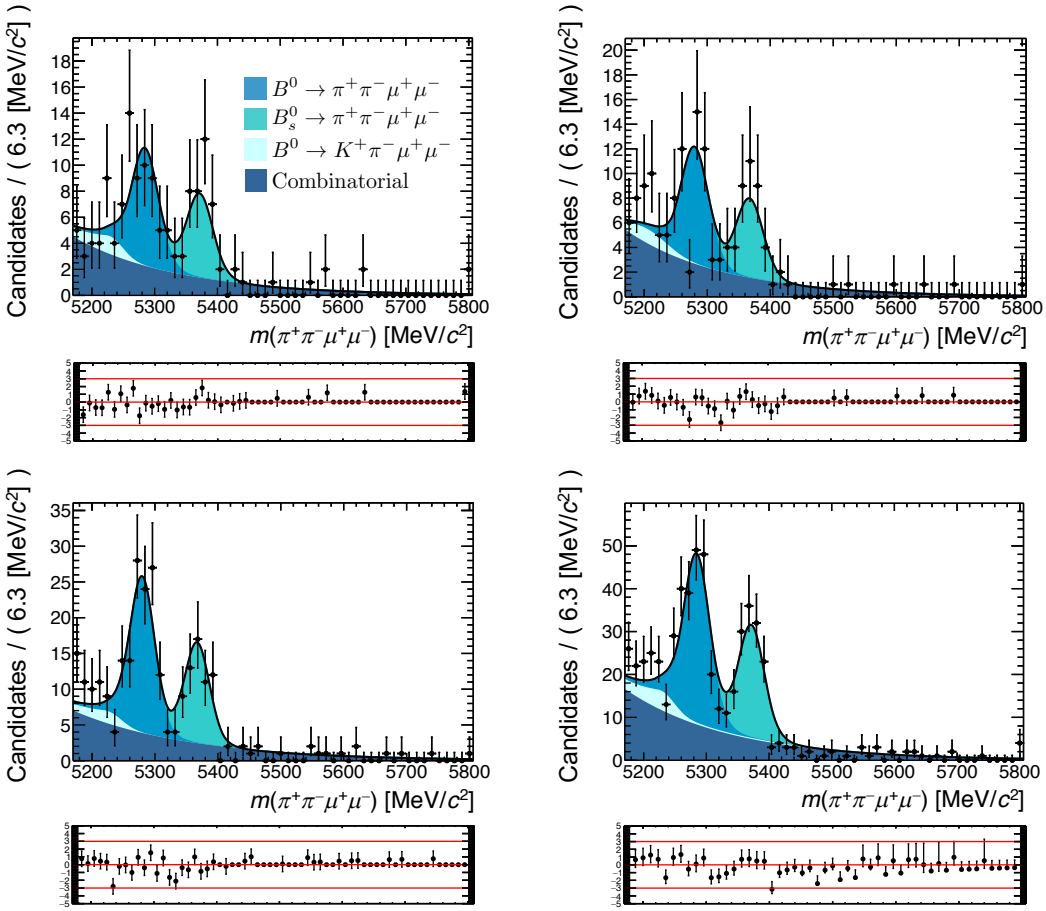

Figure C.15: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates within the full q^2 region from the data. The result of the simultaneous fit is compared to the data. The top left plot shows the fit to the 2011+2012 dataset, the top right the fit to the 2015+2016 dataset and the bottom left the fit to the 2017+2018 dataset. The bottom right plot combines the datasets and the fits for visualisation.

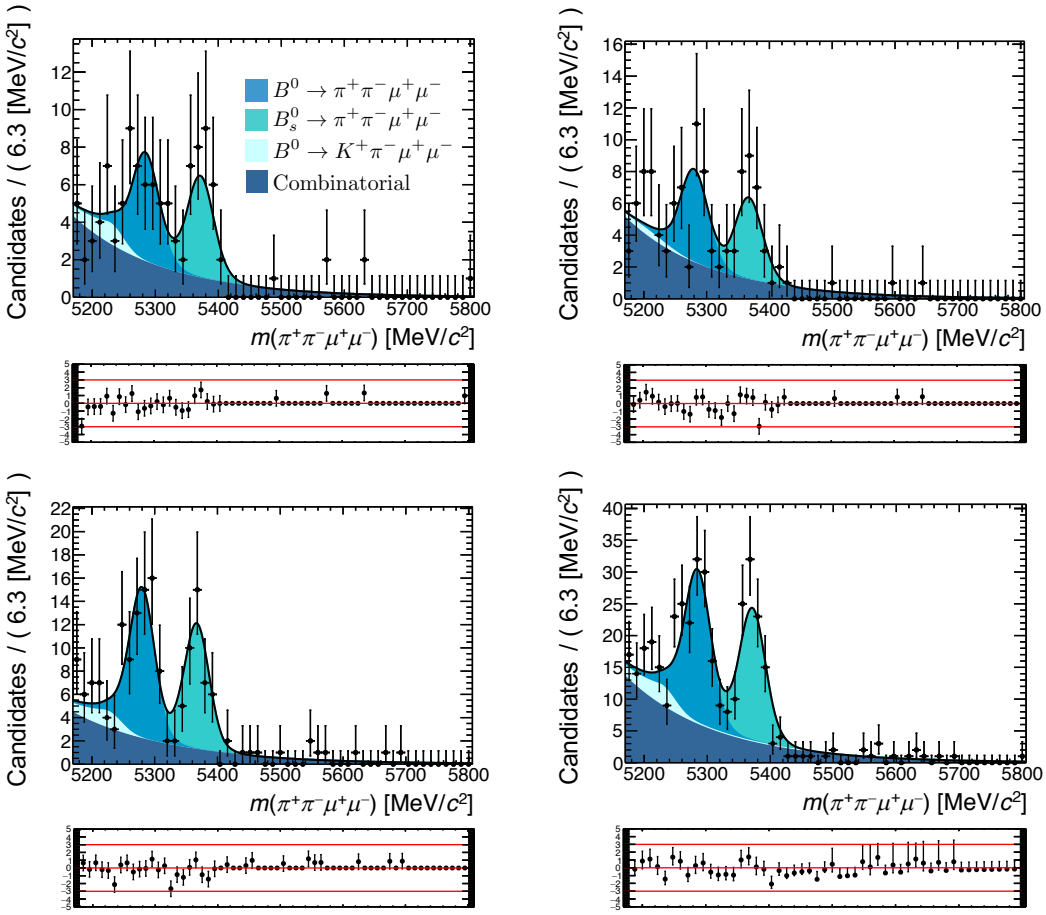

Figure C.16: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates within the low q^2 region from the data. The result of the simultaneous fit is compared to the data. The top left plot shows the fit to the 2011+2012 dataset, the top right the fit to the 2015+2016 dataset and the bottom left the fit to the 2017+2018 dataset. The bottom right plot combines the datasets and the fits for visualisation.

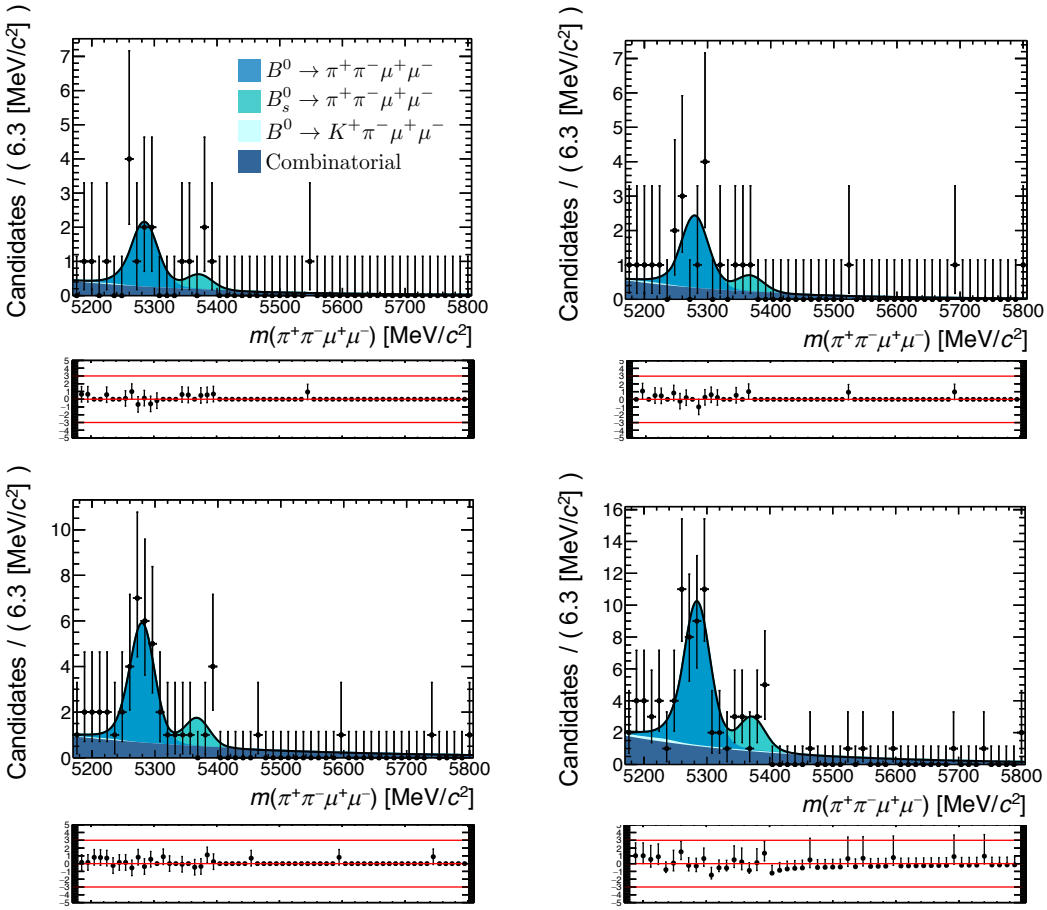

Figure C.17: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates within the high q^2 region from the data. The result of the simultaneous fit is compared to the data. The top left plot shows the fit to the 2011+2012 dataset, the top right the fit to the 2015+2016 dataset and the bottom left the fit to the 2017+2018 dataset. The bottom right plot combines the datasets and the fits for visualisation.

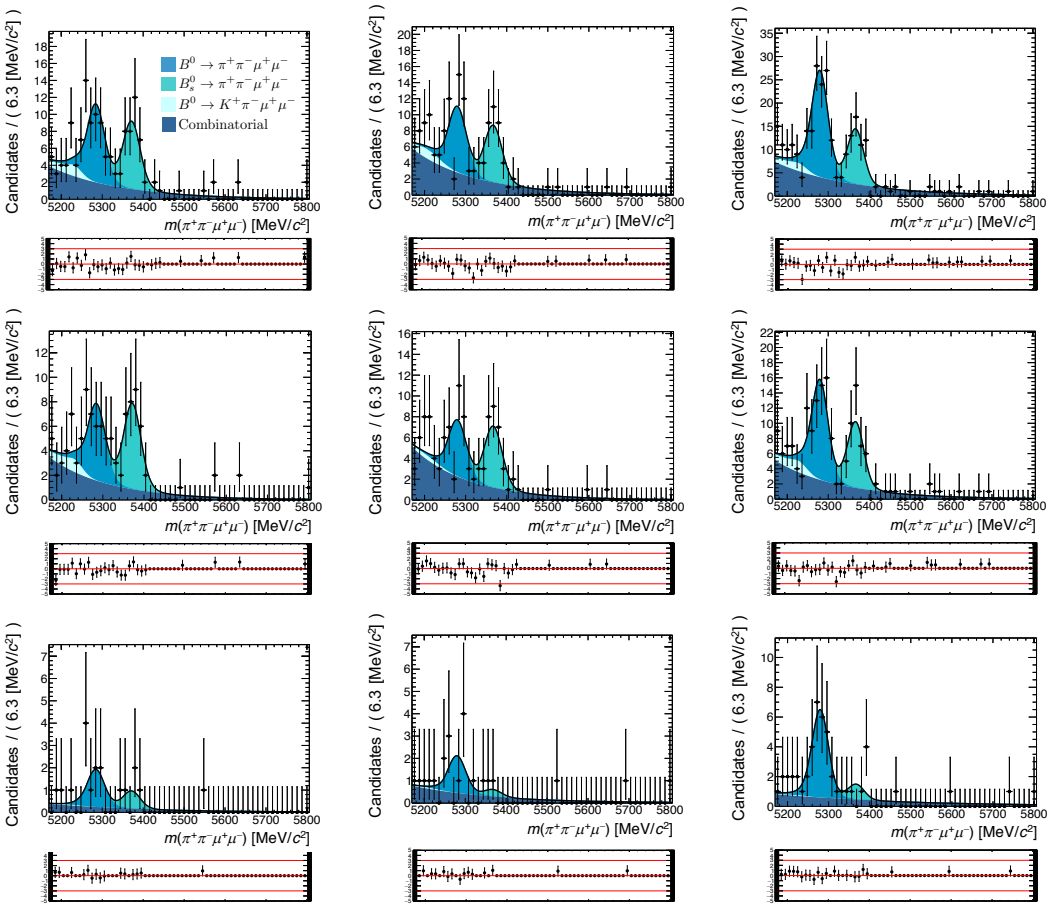

Figure C.18: The invariant mass of the $\pi^+\pi^-\mu^+\mu^-$ candidates from the data compared to the invariant mass fit when fitting each dataset separately. The left column shows the fits to the 2011+2012 dataset, the middle column the fits to the 2015+2016 dataset and the right column the fit to the 2017+2018 dataset. The top row shows the results for the full q^2 region, the middle row the results for the low q^2 region and the bottom row the results for the high q^2 region.

Bibliography

[1] H. Von Glasenapp, *Jainism: An Indian religion of salvation*, Motilal Banarsidass Publishers, (1999).

[2] A. G. Van Melsen, *From atomos to atom: The history of the concept atom*, Courier Corporation, (2004).

[3] A. Dhanani, in *Atomism in Islamic Thought*, Springer Netherlands, (2016).

[4] R. Kargon, *Walter Charleton, Robert Boyle, and the acceptance of Epicurean atomism in England*, ISIS **55** (1964) 184.

[5] R. E. Schofield, *Atomism from Newton to Dalton*, Am. J. Phys. **49** (1981) 211.

[6] J. Z. Buchwald and A. Warwick, *Histories of the electron: the birth of microphysics*, MIT Press, (2004).

[7] J. L. Heilbron, *Ernest Rutherford: And the Explosion of Atoms*, Oxford University Press, (2003).

[8] J. Chadwick, *The existence of a neutron*, Proceedings of the Royal Society of London. Series A **136** (1932) 692.

[9] M. Jammer *et al.*, *The conceptual development of quantum mechanics*, McGraw-Hill New York, (1966).

[10] J. C. Street and E. Stevenson, *New evidence for the existence of a particle of mass intermediate between the proton and electron*, Phys. Rev. **52** (1937) 1003.

[11] L. Hoddeson *et al.*, *The rise of the standard model: A history of particle physics from 1964 to 1979*, Cambridge University Press, (1997).

[12] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[13] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[14] Wikimedia Commons, *Standard model of elementary particles*, https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. Accessed: 2019-09-09, Public domain.

[15] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[16] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.

[17] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[18] F. Halzen and A. D. Martin, *Quark & Leptons: An Introductory Course In Modern Particle Physics*, John Wiley & Sons, (2008).

[19] I. J. Aitchison and A. J. Hey, *Gauge theories in particle physics: A practical introduction, Volume 1: From Relativistic Quantum Mechanics to QED*, CRC Press, (2012).

[20] S. J. Blundell and K. M. Blundell, *Concepts in thermal physics*, OUP Oxford, (2009).

[21] M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.

[22] D. Griffiths, *Introduction to elementary particles*, John Wiley & Sons, (2008).

[23] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014) 222002, arXiv:1404.1903.

[24] Belle collaboration, R. Mizuk *et al.*, *Dalitz analysis of $B^0 \to K\pi^+\psi'$ decays and the Z(4430)+*, Phys. Rev. **D80** (2009) 031104, arXiv:0905.2869.

[25] BESIII collaboration, M. Ablikim *et al.*, *Observation of a Charged Charmoniumlike Structure in $e^+ + e^- \to \pi^+\pi^- J/\psi$ at $\sqrt{s} = 4.25\,\text{GeV}$*, Phys. Rev. Lett. **110** (2013) 252001, arXiv:1303.5949.

[26] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118** (2017) 022003, arXiv:1606.07895.

[27] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[28] M.-L. Du *et al.*, *Evidence that the LHCb P_c states are hadronic molecules and the existence of a narrow $P_c(4380)$*, arXiv:1910.11846.

[29] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$*, Phys. Rev. Lett. **122** (2019) 222001, arXiv:1904.03947.

[30] S.-K. Choi *et al.*, *Observation of a narrow charmoniumlike state in exclusive $B^{\pm} \to K^{\pm}\pi^+\pi^- J/\psi$*, Physical review letters **91** (2003) 262001.

[31] S.-K. Choi *et al.*, *Observation of a resonancelike structure in the $\pi^{\pm}$ ψ' mass distribution in exclusive $B \to K^{\pm}\psi'$ decays*, Physical review letters **100** (2008) 142001.

[32] M. Gell-Mann, *A schematic model of baryons and mesons*, in *Murray Gell-Mann: Selected Papers*, pp. 151–152. World Scientific, (2010).

[33] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 1 and 2*, Developments in the quark theory of hadrons, volume 1. Edited by D. Lichtenberg and S. Rosen. pp. 22-101 (1964) 22.

[34] C. Rovelli, *Notes for a brief history of quantum gravity*, arXiv preprint gr-qc/0006061 (2000).

[35] D. Hanneke, S. Fogwell, and G. Gabrielse, *New Measurement of the Electron Magnetic Moment and the Fine Structure Constant*, Phys. Rev. Lett. **100** (2008) 120801, arXiv:0801.1134.

[36] Y. Sofue and V. Rubin, *Rotation curves of spiral galaxies*, Annu. Rev. Astron. Astr. **39** (2001) 137.

[37] R. Massey, T. Kitching, and J. Richard, *The dark matter of gravitational lensing*, Rep. Prog. Phys. **73** (2010) 086901.

[38] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, JETP lett. **5** (1967) 24.

[39] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe*, New J. Phys. **14** (2012) 095012, arXiv:1204.4186.

[40] C. Baker *et al.*, *Improved experimental limit on the electric dipole moment of the neutron*, Phys. Rev. Lett. **97** (2006) 131801.

[41] C. Csáki and P. Tanedo, *Beyond the standard model*, arXiv:1602. 04228 (2016).

[42] A. Ereditato, T. Kajita, and A. Masiero, *Focus on particle physics at the TeV scale*, New. J. Phys. **9** (2007) .

[43] E. Noether, *Invariant variation problems*, Transport Theor. Stat. **1** (1971) 186.

[44] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, CRC Press, (2018).

[45] S. Moretti, *A simple symmetry as a guide toward new physics beyond the standard model*, Front Phys. **1** (2013) 10.

[46] G. TE Velasco, *Model building and phenomenology in grand unified theories*, (2015).

[47] C.-S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[48] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2 π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.

[49] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral b meson system*, Phys. Rev. Lett. **87** (2001) 091802.

[50] BaBar collaboration, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801.

[51] LHCb collaboration, R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110** (2013) 221601.

[52] LHCb collaboration, R. Aaij *et al.*, *Observation of cp violation in charm decays*, Phys. Rev. Lett. **122** (2019) 211803.

[53] LHCb collaboration, R. Aaij *et al.*, *Observations of $\Lambda_b^0 \to \Lambda K^+\pi^-$ and $\Lambda_b^0 \to \Lambda K^+K^-$ decays and searches for other Λ_b^0 and Ξ_b^0 decays to Λh^+h^- final states*, JHEP **05** (2016) 081, arXiv:1603.00413.

[54] LHCb collaboration, R. Aaij *et al.*, *Measurement of matter-antimatter differences in beauty baryon decays*, Nature Physics **13** (2017) 391, arXiv:1609.05216.

[55] LHCb collaboration, R. Aaij *et al.*, *Observation of the decay $\Lambda_b^0 \to pK^-\mu^+\mu^-$ and search for CP violation*, JHEP **06** (2017) 108, arXiv:1703.00256.

[56] T2K collaboration, A. Dabrowska, *Results and perspectives from T2K on CPV in the neutrino sector*, Nuovo Cimento **C40** (2017) 161.

[57] J. Lees *et al.*, *Observation of time-reversal violation in the B^0 meson system*, Phys. Rev. Lett. **109** (2012) 211801.

[58] G. Lüders, *Proof of the TCP theorem*, Ann. Phys. New York **2** (1957) 1.

[59] W. Pauli, *Exclusion principle, lorentz group and reflection of space-time and charge*, in *Wolfgang Pauli*, pp. 459–479. Springer, (1988).

[60] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[61] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[62] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[63] G. S. Guralnik, C. R. Hagen, and T. W. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[64] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *A historical profile of the higgs boson*, Eur. Phys. J. C (2012) .

[65] Wikimedia Commons, *Standard model of elementary particles–most complete diagram*, `https://commons.wikimedia.org/wiki/File:Standard_Model_Of_Particle_Physics--Most_Complete_Diagram.png`. Accessed: 2019-09-16, Public domain.

[66] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Progress of Theoretical Physics **49** (1973) 652.

[67] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[68] J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. C **41** (2005) 1.

[69] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

[70] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[71] HyperCP (E871) collaboration, D. Ambrose *et al.*, *Improved branching ratio measurement for the decay $K^0_{\rm L} \to \mu^+\mu^-$*, Phys. Rev. Lett. **84** (2000) 1389.

[72] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the LHC*, Annu. Rev. Nucl. Part. S. **65** (2015) 113–143.

[73] J. L. Lopez, *Supersymmetry: From the Fermi scale to the Planck scale*, Rept. Prog. Phys. **59** (1996) 819, arXiv:hep-ph/9601208.

[74] T. Appelquist and J. Carazzone, *Infrared Singularities and Massive Fields*, Phys. Rev. **D11** (1975) 2856.

[75] F. L. Wilson, *Fermi's theory of beta decay*, Am. J. Phys. **36** (1968) 1150.

[76] E. Fermi, *An attempt of a theory of beta radiation. 1.*, Z. Phys. **88** (1934) 161.

[77] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125–1244.

[78] K. Chetyrkin, M. Misiak, and M. Münz, *Weak radiative b-meson decay beyond leading logarithms*, Physics Letters B **400** (1997) 206–219.

[79] K. G. Wilson, *Nonlagrangian models of current algebra*, Phys. Rev. **179** (1969) 1499.

[80] E. Lunghi and J. Matias, *Huge right-handed current effects in $B^0 \to K^*(K\pi)l^+l^-$ in supersymmetry*, JHEP **04** (2007) 058, arXiv:hep-ph/0612166.

[81] G. Eilam, J. L. Hewett, and A. Soni, *Rare decays of the top quark in the standard and two-higgs-doublet models*, Phys. Rev. D **44** (1991) 1473.

[82] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V l^+ l^-$ in the standard model from light-cone sum rules*, J. High. Energy Phys. **2016** (2016) 98.

[83] T. Blake, G. Lanfranchi, and D. M. Straub, *Rare B Decays as Tests of the Standard Model*, Prog. Part. Nucl. Phys. **92** (2017) 50, arXiv:1606.00916.

[84] F. Kruger and L. M. Sehgal, *CP violation in the exclusive decays $B^0 \to \pi e^+ e^-$ and $B^0 \to \rho^0 e^+ e^-$*, Phys. Rev. **D56** (1997) 5452, arXiv:hep-ph/9706247, [Erratum: Phys. Rev.D 60,099905(1999)].

[85] D. Melikhov, N. Nikitin, and S. Simula, *Rare exclusive semileptonic $b \to s$ transitions in the standard model*, Phys. Rev. **D57** (1998) 6814, arXiv:hep-ph/9711362.

[86] M. Beneke, T. Feldmann, and D. Seidel, *Exclusive radiative and electroweak $b \to d$ and $b \to s$ penguin decays at NLO*, Eur. Phys. J. **C41** (2005) 173, arXiv:hep-ph/0412400.

[87] V. M. Braun, *Light cone sum rules*, in *4th International Workshop on Progress in Heavy Quark Physics*, pp. 105–118, 9, 1997, arXiv:hep-ph/9801222.

[88] P. Ball and R. Zwicky, *New results on $B^0 \to \pi, K, \eta$ decay form factors from light-cone sum rules*, Phys. Rev. D **71** (2005) 014015.

[89] C. T. H. Davies *et al.*, *High-precision lattice QCD confronts experiment*, Phys. Rev. Lett. **92** (2004) .

[90] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (1974) 2445.

[91] LHCb collaboration, R. Aaij *et al.*, *Measurement of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction*, JHEP **11** (2016) 047, Erratum ibid. **04** (2017) 142, arXiv:1606.04731.

[92] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, arXiv:1506.08777.

[93] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^*\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044.

[94] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of $B^0 \to K^{*0}\mu^+\mu^-$ at large recoil*, J. High Energy Phys. **2013** (2013) 48.

[95] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.

[96] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **10** (2018) 047, arXiv:1805.04000.

[97] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B781** (2018) 517, arXiv:1710.02846.

[98] Belle collaboration, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^* \ell^+ \ell^-$*, Phys. Rev. Lett. **118** (2017) 111801, arXiv:1612.05014.

[99] W. Altmannshofer and D. M. Straub, *New physics in $b \to s$ transitions after LHC run 1*, Eur. Phys. J. **C75** (2015) 382, arXiv:1411.3161.

[100] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *On the impact of power corrections in the prediction of $B \to K^* \mu^+ \mu^-$ observables*, JHEP **12** (2014) 125, arXiv:1407.8526.

[101] J. Lyon and R. Zwicky, *Resonances gone topsy turvy-the charm of QCD or new physics in $b \to s l^+ l^-$*, arXiv:1406. 0566 (2014).

[102] M. Ciuchini *et al.*, *$B^0 \to K^{*0} l^+ l^-$ decays at large recoil in the standard model: a theoretical reappraisal*, J. High Energy Phys. **2016** (2016) 116.

[103] V. Chobanova *et al.*, *Large hadronic power corrections or new physics in the rare decay $B^0 \to K^{*0} \mu^+ \mu^-$*, J. High Energy Phys. **2017** (2017) 25.

[104] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \to K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **122** (2019) 191801, arXiv:1903.09252.

[105] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[106] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0} \ell^+ \ell^-$ decays*, JHEP **08** (2017) 055, arXiv:1705.05802.

[107] BaBar collaboration, J. Lees *et al.*, *Measurement of branching fractions and rate asymmetries in the rare decays $B^0 \to K^{*0} l^+ l^-$*, Phys. Rev. D **86** (2012) 032012.

[108] Belle collaboration, J.-T. Wei *et al.*, *Measurement of the differential branching fraction and forward-backward asymmetry for $B^0 \to K^{*0} l^+ l^-$*, Phys. Rev. Lett. **103** (2009) 171801.

[109] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\overline{B}^0 \to D^{*+} \tau^- \overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+} \mu^- \overline{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, Publisher's Note ibid. **115** (2015) 159901, arXiv:1506.08614.

[110] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $\mathcal{B}(B^0 \to D^{*-} \tau^+ \nu_\tau)$ and $\mathcal{B}(B^0 \to D^{*-} \mu^+ \nu_\mu)$ branching fractions using three-prong τ-lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, arXiv:1708.08856.

[111] BaBar collaboration, J. P. Lees *et al.*, *Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, Phys. Rev. Lett. **109** (2012) 101802, arXiv:1205.5442.

[112] BaBar collaboration, J. P. Lees *et al.*, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. **D88** (2013) 072012, arXiv:1303.0571.

[113] Belle collaboration, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. **D92** (2015) 072014, arXiv:1507.03233.

[114] Belle collaboration, Y. Sato *et al.*, *Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method*, Phys. Rev. **D94** (2016) 072007, arXiv:1607.07923.

[115] Belle collaboration, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017) 211801, arXiv:1612.00529.

[116] J. Aebischer *et al.*, *B-decay discrepancies after Moriond 2019*, arXiv:1903. 10434 (2019).

[117] B. Allanach, F. S. Queiroz, A. Strumia, and S. Sun, *$Z^{'}$ models for the LHCb and $g-2$ muon anomalies*, Phys. Rev. D **93** (2016) 055045.

[118] S. Sahoo and R. Mohanta, *Scalar leptoquarks and the rare B meson decays*, Phys. Rev. D **91** (2015) 094019.

[119] M. Bauer and M. Neubert, *Minimal leptoquark explanation for the R(D*), R(K), and $(g-2)$ μ anomalies*, Phys. Rev. Lett. **116** (2016) 141802.

[120] G. W. Bennett *et al.*, *Final report of the E821 muon anomalous magnetic moment measurement at BNL*, Phys. Rev. D **73** (2006) 072003.

[121] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

[122] LHCb collaboration, R. Aaij *et al.*, *Study of the rare B_s^0 and B^0 decays into the $\pi^+\pi^-\mu^+\mu^-$ final state*, Phys. Lett. **B743** (2015) 46, arXiv:1412.6433.

[123] R. Li, C. Lü, W. Wang, and X. Wang, *$b \to s$ transition form factors in the perturbative QCD approach*, Phys. Rev. D **79** (2009) 014013.

[124] P. Colangelo, F. De Fazio, and W. Wang, *$B_s \to f_0(980)$ form factors and B_s decays into $f_0(980)$*, Phys. Rev. D **81** (2010) 074001.

[125] N. Ghahramany and R. Khosravi, *Analysis of the rare semileptonic decays of B_s to $f_0(980)$ and $K_0^*(1430)$ scalar mesons in QCD sum rules*, Phys. Rev. D **80** (2009) 016009.

[126] Z.-G. Wang and W.-M. Yang, *Analysis of the $f_0(980)$ and $a_0(980)$ mesons as four-quark states with the QCD sum rules*, Eur. Phys. J. C **42** (2005) 89, arXiv:hep-ph/0501105.

[127] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[128] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[129] L. Evans and P. Bryant, *LHC machine*, JINST **3** (2008) S08001.

[130] E. A. Mobs, *The CERN accelerator complex*, tech. rep., 2016.

[131] W. Herr and B. Muratori, *Concept of luminosity (2006)*, Yellow Report CERN **2** (2006).

[132] L. Evans and P. Bryant, *LHC machine*, J. Instrum. **3** (2008) S08001.

[133] ATLAS collaboration, G. Aad *et al.*, *The ATLAS experiment at the CERN large hadron collider*, Jinst **3** ((2008)) S08003.

[134] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, Jinst ((2008)).

[135] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, J. Instrum. **3** ((2008)) S08002.

[136] E. Norrbin and T. Sjöstrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. C **17** (2000) 137.

[137] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189.

[138] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13* TeV pp *collisions*, arXiv:1902.06794, submitted to Phys. Rev. Lett.

[139] *LHCb material for presentations*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html. Accessed: 2019-01-10.

[140] C. Abellan Beteta *et al.*, *Monitoring radiation damage in the LHCb Tracker Turicensis*, arXiv:1809.05063.

[141] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[142] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030.

[143] LHCb collaboration, R. Aaij *et al.*, *Design and performance of the lhcb trigger and full real-time reconstruction in run 2 of the lhc*, Journal of Instrumentation **14** (2019) P04013.

[144] LHCb Outer Tracker Group, P. d'Argent *et al.*, *Improved performance of the LHCb outer tracker in LHC run 2*, arXiv:1708. 00819 (2017).

[145] LHCb collaboration, A. Papanestis and C. D'Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, Nucl. Instrum. Meth. **A876** (2017) 221, arXiv:1703.08152.

[146] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[147] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031.

[148] T. Head, *The LHCb trigger system*, JINST **9** (2014) C09015. 8 p.

[149] LHCb collaboration, B. Sciascia, *LHCb Run 2 trigger performance*, PoS **BEAUTY2016** (2016) 029.

[150] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.

[151] V. V. Gligorov, M. Williams, and C. Thomas, *The HLT inclusive B triggers*, tech. rep., 2011.

[152] LHCb collaboration, R. Aaij *et al.*, *Evidence for the decay* $B_s^0 \to \overline{K}^{*0}\mu^+\mu^-$, J. High Energy Phys. **2018** (2018).

[153] Belle collaboration, K. Chilikin *et al.*Phys. Rev. D **90** (2014).

[154] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.

[155] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, arXiv:1307.2782.

[156] LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for* 7 TeV *pp collisions*, LHCb-CONF-2013-011.

[157] LHCb collaboration, R. Aaij *et al.*, *Measurement of f_s/f_u variation with proton-proton collision energy and b-meson kinematics*, Phys. Rev. Lett. **124** (2020).

[158] LHCb collaboration, R. Aaij *et al.*, *Analysis of the resonant components in $B^0 \to J/\psi \pi^+\pi^-$*, Phys. Rev. D **87** (2013) 052001.

[159] LHCb collaboration, R. Aaij *et al.*, *Analysis of the resonant components in $B_s^0 \to J/\psi \pi^+\pi^-$*, Phys. Rev. D **86** (2012) 052006.

[160] J. Albrecht, L. collaboration *et al.*, *The LHCb trigger system: Present and future*, in *Journal of Physics: Conference Series*, vol. 623, p. 012003, IOP Publishing, 2015.

[161] M. Williams *et al.*, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.

[162] F. Archilli *et al.*, *Performance of the muon identification at lhcb*, Journal of Instrumentation **8** (2013) P10020.

[163] M. De Cian, S. Stahl, P. Seyfert, and S. Farry, *Fast neural-net based fake track rejection in the lhcb reconstruction, cern public note: Cern-lhcb-pub-2017-011*, tech. rep., 2017.

[164] G. A. Cowan, D. C. Craik, and M. D. Needham, *Rapidsim: An application for the fast simulation of heavy-quark hadron decays*, Computer Physics Communications **214** (2017) 239–246.

[165] M. Cacciari, M. Greco, and P. Nason, *The $p_{\rm T}$ spectrum in heavy-flavour hadroproduction*, J. High Energy Phys. **1998** (1998) 007–007.

[166] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. A **555** (2005) 356–369.

[167] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*, CRC press, (1984).

[168] T. Chen and C. Guestrin, *XGBoost: A scalable tree boosting system*, in *Proceedings of the 22Nd ACM SIGKDD International conference on knowledge discovery and data mining*, KDD '16, p. 785, 2016, arXiv:1603.02754. doi: 10.1145/2939672.2939785.

[169] D. Derkach *et al.*, *Machine-learning-based global particle-identification algorithms at the lhcb experiment*, in *Journal of Physics: Conference Series*, vol. 1085, p. 042038, IOP Publishing, 2018.

[170] L. Anderlini *et al.*, *The PIDCalib package*, tech. rep., 2016.

[171] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

[172] LHCb collaboration, R. Aaij *et al.*, *Analysis of the resonant components in* $\overline{B}^0 \to J/\psi\pi^+\pi^-$, Phys. Rev. **D87** (2013) 052001, arXiv:1301.5347.

[173] LHCb collaboration, R. Aaij *et al.*, *Measurement of resonant and CP components in* $\overline{B}^0_s \to J/\psi\pi^+\pi^-$ *decays*, Phys. Rev. **D89** (2014) 092006, arXiv:1402.6248.

[174] LHCb collaboration, R. Aaij *et al.*, *Evidence for the decay* $B^0_s \to \overline{K}^{*0}\mu^+\mu^-$, JHEP **07** (2018) 020, arXiv:1804.07167.

[175] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[176] LHCb collaboration, R. Aaij *et al.*, *Observation of the* $\Lambda^0_b \to J/\psi p\pi^-$ *decay*, JHEP **07** (2014) 103, arXiv:1406.0755.

[177] LHCb collaboration, R. Aaij *et al.*, *Observation of the suppressed decay* $\Lambda^0_b \to p\pi^-\mu^+\mu^-$, JHEP **04** (2017) 029, arXiv:1701.08705.

[178] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation parameters and polarisation fractions in* $B^0_s \to J/\psi\,\overline{K}^{*0}$ *decays*, JHEP **11** (2015) 082, arXiv:1509.00400.

[179] L. Zhang and S. Stone, *Time-dependent dalitz-plot formalism for* $B^0_q \to J/\psi\, h^+h^-$, Phys. Lett. B **719** (2013) 383.

[180] LHCb collaboration, R. Aaij *et al.*, *Measurement of the resonant and CP components in* $\overline{B}^0 \to J/\psi\pi^+\pi^-$ *decays*, Phys. Rev. **D90** (2014) 012003, arXiv:1404.5673.

[181] E. P. Wigner, *Gruppentheorie und ihre anwendungen auf die quantenmechanik der atomspektren. friedr. vieweg und sohn akt*, Ges. Braunschweig (1931).

[182] G. Breit and E. Wigner, *Capture of slow neutrons*, Phys. Rev. **49** (1936) 519.

[183] S. M. Flatte, *Coupled-channel analysis of the* $\pi\eta$ *and kk systems near kk threshold*, Phys. Lett. B **63** (1976) 224.

[184] Belle collaboration, K. Chilikin *et al.*, *Observation of a new charged charmoniumlike state in $\overline{B} \to J/\psi K^- \pi^+$ decays*, Phys. Rev. D. **90** (2014).

[185] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, JHEP **01** (2009) 019, arXiv:0811.1214.

[186] T. Blake, U. Egede, and A. Shires, *The effect of s-wave interference on the $B^0 \to K^{*0} l^+ l^-$ angular observables*, Journal of High Energy Physics **2013** (2013).

[187] T. M. Aliev, K. Azizi, and M. Savci, *Analysis of rare $B^0 \to K^*(1430) l^+ l^-$ decay within QCD sum rules*, Phys. Rev. **D76** (2007) 074017, arXiv:0710.1508.

[188] P. Colangelo, F. De Fazio, and W. Wang, *$B_s \to f_0(980)$ form factors and B_s decays into $f_0(980)$*, Phys. Rev. **D81** (2010) 074001, arXiv:1002.2880.

[189] M. Emmerich, M. Strohmaier, and A. Schäfer, *$B \to f_2(1270)$ form factors with light-cone sum rules*, Phys. Rev. **D98** (2018) 014008, arXiv:1804.02953.

[190] K. Yang, *B to Light Tensor Meson Form Factors Derived from Light-Cone Sum Rules*, Phys. Lett. **B695** (2011) 444, arXiv:1010.2944.

[191] R. W. Johnson, *An introduction to the bootstrap*, Teaching Statistics **23** (2001) 49.

[192] LHCb collaboration, *Measurement of the track reconstruction efficiency at LHCb*, J. Instrum. **10** (2015) P02007.

[193] T. Skwarnicki, *A study of the radiative cascade transitions between the upsilon-prime and upsilon resonances*, tech. rep., DESY collaboration, (1986).

[194] R. A. Davis, K. Lii, and D. N. Politis, *Remarks on some nonparametric estimates of a density function*, in *Selected Works of Murray Rosenblatt*, pp. 95–100. Springer, 2011.

[195] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Stat. **33** (1962) 1065.

[196] I. Belyaev *et al.*, *Simulation application for the LHCb experiment*, arXiv:physics/0306035.

[197] LHCb Public Archive, *Large hadron collider beauty*, https://public-archive.web.cern.ch/en/LHC/LHCb-en.html. Accessed: 2020-03-02, Public domain.

[198] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867.

[199] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[200] G. Milluzzo *et al.*, *Geant4 simulation of the ELIMED transport and dosimetry beam line for high-energy laser-driven ion beam multidisciplinary applications*, Nucl. Instrum. Meth. A **909** (2018) 298, arXiv:1802.03745.

[201] *LHCb Boole web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[202] *LHCb Brunel web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/.

[203] *LHCb Moore web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/moore/.

[204] *LHCb DaVinci web page*, http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[205] Corti, G. and others, *Monte carlo event type definition rules*, LHCb-2005-034.